UNIVERSITY OF LONDON THESIS

Degree PHD Year 2008 Name of Author BOKHOVEN, MARIEKE, CHRISTINA.

Measurement of insertional mutagenesis by retroviral and lentiviral vectors

Marieke Christina Bokhoven

Submitted to the University of London for the degree of

Doctor of Philosophy

May 2008

Division of Infection and Immunity

Windeyer Building

University College London

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Declaration

I, Marieke Christina Bokhoven, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Retroviral vectors have been successfully used in the clinic for the correction of inherited immunodeficiencies. However, in 2 recent gene therapy trials for X-linked Severe Combined Immunodeficiency, 5 patients developed leukaemia. This was associated with vector integration near cellular proto-oncogenes, leading to their activation; a process known as insertional mutagenesis.

This thesis describes the development of a cell line assay that allows us to quantify insertional mutagenesis by retroviral and lentiviral vectors and to analyse the mechanisms by which this occurs. The interleukin-3 (IL-3) dependent cell line BCL15 is derived from the mouse bone marrow cell line BAF3 and over-expresses human Bcl2. The frequency at which IL-3 independent mutants are obtained following vector transduction is measured. Lentiviral and retroviral vectors transform BCL15 cells at similar integrant frequencies of 4.3×10^{-8} and 1.2×10^{-7}, respectively. However, they cause insertional mutagenesis in this assay by different mechanisms. The human immunodeficiency virus-1 (HIV-1)-derived lentiviral vector HV transforms BCL15 cells by insertional activation of the growth hormone receptor (Ghr) gene. An HIV-Ghr fusion transcript was detected. It originates from the HIV-1 5'-LTR; it then splices from the HIV-1 major splice donor to the splice acceptor of Ghr exon 2. The mutants express GHR and grow in response to bovine growth hormone in the foetal calf serum of the culture medium. Deletion of the HIV-1 enhancer/promoter in a self-inactivating vector prevents this mechanism of transformation. Retroviral vector transformation of the BCL15 cell line does not occur via Ghr gene activation. Retroviral vectors up-regulate expression of the cytokine IL-3, either by insertion into the IL-3 gene or by insertion into other genes, which may act as upstream activators of IL-3 expression. This assay is a general method to quantitate insertional mutagenesis. It may inform the design of safer vectors and can be used in initial safety testing of (pre-) clinical vectors.

Acknowledgements

First and foremost I would like to thank my supervisors Yasu Takeuchi and Mary Collins for allowing me to pursue my research interests in their laboratories. It has been a wonderful opportunity and I am very grateful for all their help, advice and words of encouragement over the past four years.

It was a privilege to have been part of both the Collins laboratory and the Wohl Virion Centre during my PhD. I thank everyone for their friendly help and for creating such a stimulating research environment to work in. In particular, I would like to thank David Escors for providing invaluable practical help throughout. I thank my office colleagues Tim Harrison and Nigel Temperton for their advice and particularly IT assistance. I thank all Windeyer support staff, especially John Cheshire, for facilitating my work in countless ways.

I would like to thank all past and present members of "Team IM". A special thank you to Sam Stephen for the many hours he spent training me at the start of my PhD. It was a pleasure working with Sean Knight over the past few months and I wish him well in continuing this project. I thank Conrad Vink for help with setting up LM-PCR and Evelien Gevers for performing phospho-STAT5 staining.

I met some wonderful friends at the Windeyer, who made my work in the lab and life outside of it so much more enjoyable. Thank you so much to Luis Apolonia, Ranbir Bailey, Sofia Efklidou, Sabrina Ivol, Sean Knight, Giada Mattiuzzo, Helen Rowe and Emma Willoughby. I especially thank Ranbir and Sean for proofreading parts of this thesis.

I would like to thank our MB PhD programme tutor, Gordon Stewart for all his valuable advice over the years. His early morning teaching sessions on the wards, followed by coffee, were always a welcome break from lab work.

I am very grateful to the International Journal of Experimental Pathology and the Department of Health for generously funding my PhD studentship.

Lastly, I would like to thank my family. I thank my grandfather Cees Bokhoven for encouraging my interest in science. Unfortunately, he did not live to see the completion of this thesis. I am eternally grateful to my parents and my sister for all their love and support. To them I dedicate this thesis.

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

List of Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Abbreviations

[3H]	tritium
3'UTR	3' untranslated region
6TG	6-thioguanine
AAV	adeno-associated virus
ADA-SCID	adenosine-deaminase severe combined immunodeficiency
AEV	avian erythroblastosis virus
AIDS	acquired immunodeficiency syndrome
ALL	acute lymphocytic leukaemia
ALV	avian leukosis virus
AML	acute myeloid leukaemia
A-MLV	Abelson murine leukaemia virus
AP-1	activator protein 1
AP-2	activator protein 2
APC	allophycocyanin
APOBEC	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like
ARE	AU-rich element
ASLV	avian sarcoma and leukosis virus
ATP	adenosine triphosphate
BAF	barrier to autointegration factor
bGH	bovine growth hormone
bHLH	basic helix loop helix
BM	bone marrow
C/EBP	CCAAT/enhancer binding protein
CA	capsid
CAT-1	cationic amino acid transporter 1
CBF	core binding factor
CCR5	chemokine (C-C motif) receptor 5
CDK	cyclin dependent kinase
CGD	chronic granulomatous disease
CIS	common insertion site
CMV	cytomegalovirus
CNS	central nervous system
c-onc	cellular oncogene
cPPT	central polypurine tract
CTD	carboxy terminal domain
CTE	constitutive transport element
CXCR4	chemokine (C-X-C motif) receptor 4
DC	dendritic cells
DEPC	diethylpyrocarbonate
dLNGFR	deleted form of the low-affinity nerve growth factor receptor
DMEM	Dulbecco's Modified Eagle's Medium
DNA	deoxyribonucleic acid

dNTPs	deoxynucleotide triphosphates
ds	double stranded
E.coli	Escherichia coli
EBV	Epstein Barr virus
EF1α	elongation factor 1 alpha
EGF	epidermal growth factor
EGFR	epidermal growth factor receptor
EIAV	equine infectious anaemia virus
EPO	erythropoietin
EPO-R	erythropoietin receptor
ERK	Extracellular signal Regulated Kinases
ESCRT	endosomal sorting complex required for transport
FCS	foetal calf serum
FeLV	feline leukaemia virus
FGF	fibroblast growth factor
FGFR	fibroblast growth factor receptor
FIV	feline immunodeficiency virus
Fr-MCF	Friend mink cell focus-forming
Fr-MLV	Friend murine leukaemia virus
GALT	gut-associated lymphoid tissue
GALV	gibbon ape leukaemia virus
G-CSF	granulocyte-colony stimulating factor
G-CSF-R	granulocyte-colony stimulating factor receptor
GDP	guanosine diphosphate
GFP	green fluorescent protein
GH	growth hormone
GHR	growth hormone receptor
GM-CSF	granulocyte-macrophage colony-stimulating factor
GM-CSF-R	granulocyte-macrophage colony-stimulating factor receptor
GR	glucocorticoid receptor
GRE	glucocorticoid response element
GTP	guanosine triphosphate
h	hours
HAART	highly active anti-retroviral therapy
HAT	histone acetyltransferase
HBSS	Hanks' Balanced Salt Solution
HCC	hepatocellular carcinoma
HDAC	histone deacetylase
HHV	human herpes virus
HIV	human immunodeficiency virus
HLA	human leukocyte antigen
HMGA-1	high mobility group AT-hook 1
hnRNP	heterogeneous nuclear ribonucleoprotein
Hprt	hypoxanthine-guanine phosphoribosyltransferase
HPV	human papilloma virus
HSC	haematopoietic stem cell

HSCT	haematopoietic stem cell transplantation
HTLV	Human T cell leukaemia virus
IAP	intracisternal A particle
IGF-1	insulin-like growth factor 1
IL-#	interleukin-number
IL-#-R	interleukin-number-receptor
IN	integrase
Ini-1	Integrase interactor 1
INS	inhibitory sequences
IPCR	inverse PCR
IRES	internal ribosome entry site
IκB	inhibitor of kappa B
JAK	Janus kinase
KS	Kaposi sarcoma
KSHV	Kaposi sarcoma herpes virus
LB	Luria-Bertani
LEDGF	lens epithelium-derived growth factor
LEF-1	Lymphoid enhancer-binding factor 1
LEM	(Lap, emerin, MAN)
LM-PCR	linker-mediated PCR
LTR	long terminal repeat
MA	matrix
MAPK	mitogen-activated protein kinases
MCF	mink cell focus-forming virus
MCREF-1	mammalian type C retrovirus enhancer factor 1
MDR	multidrug resistance
min	minutes
miRNA	micro-ribonucleic acid
MLV	murine leukaemia virus
MLV-A	amphotropic murine leukaemia virus
MLV-E	ecotropic murine leukaemia virus
MLV-P	polytropic murine leukaemia virus
MLV-X	xenotropic murine leukaemia virus
MMTV	mouse mammary tumour virus
MOI	multiplicity of infection
Mo-MCF	Moloney mink cell focus-forming
Mo-MLV	Moloney murine leukaemia virus
MPSV	myeloproliferative sarcoma virus
mRNA	messenger ribonucleic acid
MSCV	murine stem cell virus
MSD	major splice donor
NC	nucleocapsid
NF-1	nuclear factor 1
NFAT	nuclear factor of activated T-cells
NFκB	nuclear factor-kappa B
NHL	non-Hodgkin's lymphoma

NLS	nuclear localisation signal
NPC	nuclear pore complex
ORF	open reading frame
PB	polybrene
PBL	peripheral blood lymphocytes
PBS	primer binding site
PBS	phosphate buffered saline
PCP	pneumocystis pneumonia
PCR	polymerase chain reaction
PDGF	platelet derived growth factor
PDGFR	platelet derived growth factor receptor
PEG-ADA	polyethylene glycol adenosine deaminase
PEL	primary effusion lymphoma
PGK	phosphoglycerate kinase
PI	propidium iodide
PI3K	phosphoinositide-3 kinase
PIC	pre-integration complex
PIP3	phosphatidylinositol (3,4,5)-trisphosphate
Pit-2	inorganic phosphate transporter 2
PMA	phorbol myristate acetate
PPT	polypurine tract
PR	protease
PRL	prolactin
PRL-R	prolactin receptor
pTEF-b	positive transcription elongation factor-b
Q-PCR	quantitative PCR
Q-RT-PCR	Quantitative reverse transcriptase PCR
R	repeat region
RARE	retinoic acid response element
RCR	replication competent retrovirus
RD114	feline endogenous retrovirus
RER	rough endoplasmatic reticulum
RIS	retroviral integration site
RNA	ribonucleic acid
RNAPII	ribonucleic acid polymerase II
RRE	Rev-response element
RSV	Rouse sarcoma virus
RT	reverse transcriptase
RTC	reverse transcription complex
RTCGD	retrovirus tagged cancer gene database
RT-PCR	reverse transcriptase PCR
SA	splice acceptor
SCF	stem cell factor
scFc	single-chain antibody
SD	splice donor
SFFV	spleen focus forming virus

SIN	self-inactivating
SIV	simian immunodeficiency virus
snoRNA	small nucleolar ribonucleic acid
snRNP	small nuclear ribonucleoproteins
ss	single-stranded
STAT	Signal Transducers and Activators of Transcription protein
SU	surface unit
TAR	trans-activation responsive element
TFIIH	Transcription Factor II H
TFT	trifluorothymidine
TIL	tumour infiltrating lymphocytes
Tk	thymidine kinase
TM	transmembrane
TPO	thrombopoietin
TRIM	TRIpartite Motif
tRNA	transfer ribonucleic acid
TU	transcription unit
U3	unique 3' region
U5	unique 5' region
USE	upstream sequence element
USF-1	upstream stimulatory factor 1
VLP	virus-like particle
v-onc	viral oncogene
VSV-G	vesicular stomatis virus G protein
WAS	Wiskott-Aldrich syndrome
WPRE	woodchuck hepatitis virus post-transcriptional regulatory element
WT	wild type
X-SCID	X-linked severe combined immunodeficiency
βc	common beta chain
βIL-3	interleukin-3 receptor beta chain
γc	common gamma chain
Ψ	psi packaging signal

Chapter 1

1. Introduction

1.1. Retrovirus Biology

Retroviruses (*Retroviridae* family) are a unique class of viruses that have single stranded (ss) RNA genomes, which serve as templates for reverse transcription into double stranded (ds) DNA by an RNA-dependent DNA polymerase known as reverse transcriptase. A further distinctive feature of retrovirus replication is, that this reverse transcribed dsDNA genome becomes integrated into the host cell genome. This property makes vectors derived from these viruses particularly suitable for use in gene therapy.

In this thesis, vectors derived from the genomes of Murine Leukaemia Virus (MLV) and Human Immunodeficiency Virus-1 (HIV-1) were used. In the ensuing text I shall therefore focus on distinctive features of these two viruses. As we believe these may play a role in vector-mediated pathogenesis, two stages of the retroviral life cycle will receive particular attention: integration and long terminal repeat (LTR)-driven gene expression.

1.1.1. Retrovirus Taxonomy

The *retroviridae* family can be subdivided into seven genera, based on both genome structure (See section 1.1.2.) and virion morphology (Table 1.1). MLV is a gammaretrovirus; HIV-1 is a lentivirus.

1.1.2. Retroviral genome structure

The RNA genome structures of Moloney MLV (Mo-MLV) and HIV-1 are depicted in Figure 1.1. They are 8332bp and 9181nt in size, respectively.

Gene regulatory sequences are found at the 5' and 3' ends of all retroviral genomes. In the RNA genome, at the 5' end, R (repeat) and U5 (unique 5) segments are found; the 3'end is made up of a U3 (unique 3) and an identical R segment. Simple retroviruses such as Mo-MLV contain 3 genes: *gag*, *pol* and *env*. The MLV *gag* gene encodes the 3 viral structural proteins: matrix (MA), capsid (CA) and nucleocapsid (NC) and the polypeptide p12. The *pol* gene encodes the viral enzymes necessary for replication: protease (PR), reverse transcriptase (RT)/ RNase H and integrase (IN).

The *env* gene encodes transmembrane (TM) and surface (SU) glycoproteins. They are found in the viral envelope, a lipid bilayer derived from the cell membrane, and mediate virus entry into a cell by interacting with specific receptors on a cell surface. HIV-1 is a complex retrovirus. Its genome contains an additional 6 genes: the regulatory genes *tat* and *rev*; and the accessory genes *nef*, *vpu*, *vpr* and *vif*. The functions of these genes will be discussed. In addition to MA, CA and NC, HIV-1 *gag* encodes the polypeptide p6.

1.1.3. Retroviral virion structure

Figure 1.2. depicts the structure of an HIV-1 virion. Two copies of genomic RNA with associated primer tRNAs (tRNA(Pro) for MLV (Harada *et al.* 1979); tRNA(3Lys) for HIV-1 (Jiang *et al.* 1993)) are complexed with nucleocapsid protein. This RNA/protein complex is packaged together with the viral enzymes (PR, RT, IN) and cellular proteins in a protein shell made up of capsid protein. MLV has a roughly spherical core, whereas the HIV-1 core is elongated and conical. Matrix protein is found between the capsid and the lipid membrane, in which envelope glycoproteins are embedded.

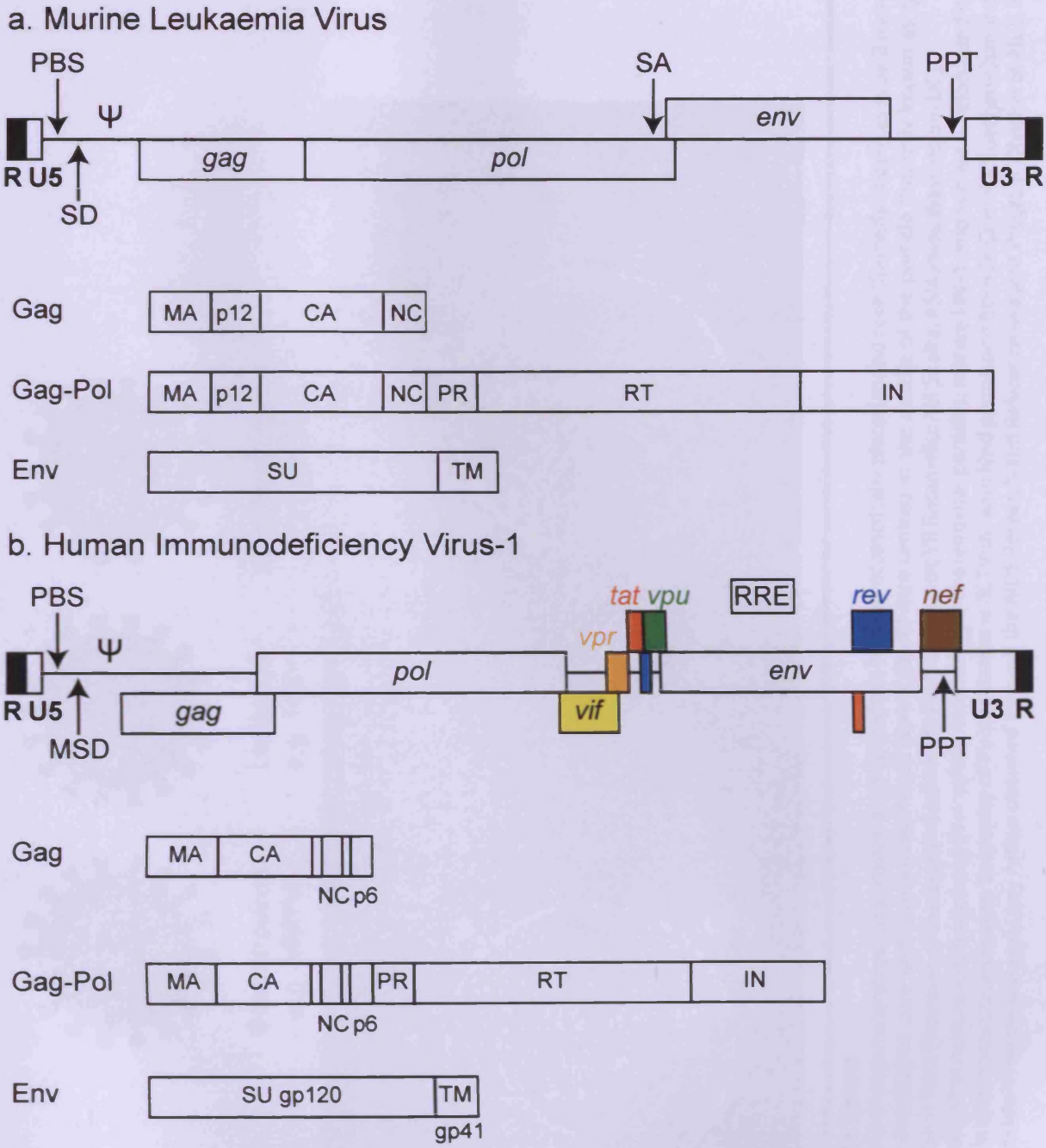


Figure 1.1. MLV and HIV-1 genome organisation
The RNA genome structures of MLV (a) and HIV-1 (b) are shown. The gag, pol and env genes are common to all retroviruses. HIV-1 regulatory genes tat and rev and accessory genes vif, vpr, vpu and nef are located downstream of pol. They partially overlap with env, U3 or each other. Tat and rev are each made up of 2 exons which are found in different reading frames. The organisation of Gag, Gag-Pol and Env polyprotein precursors is shown below each genome structure.

a.

Betaretrovirus
e.g. MMTV

Lentivirus
e.g. HIV-1

Gammaretrovirus
e.g. MLV

b.

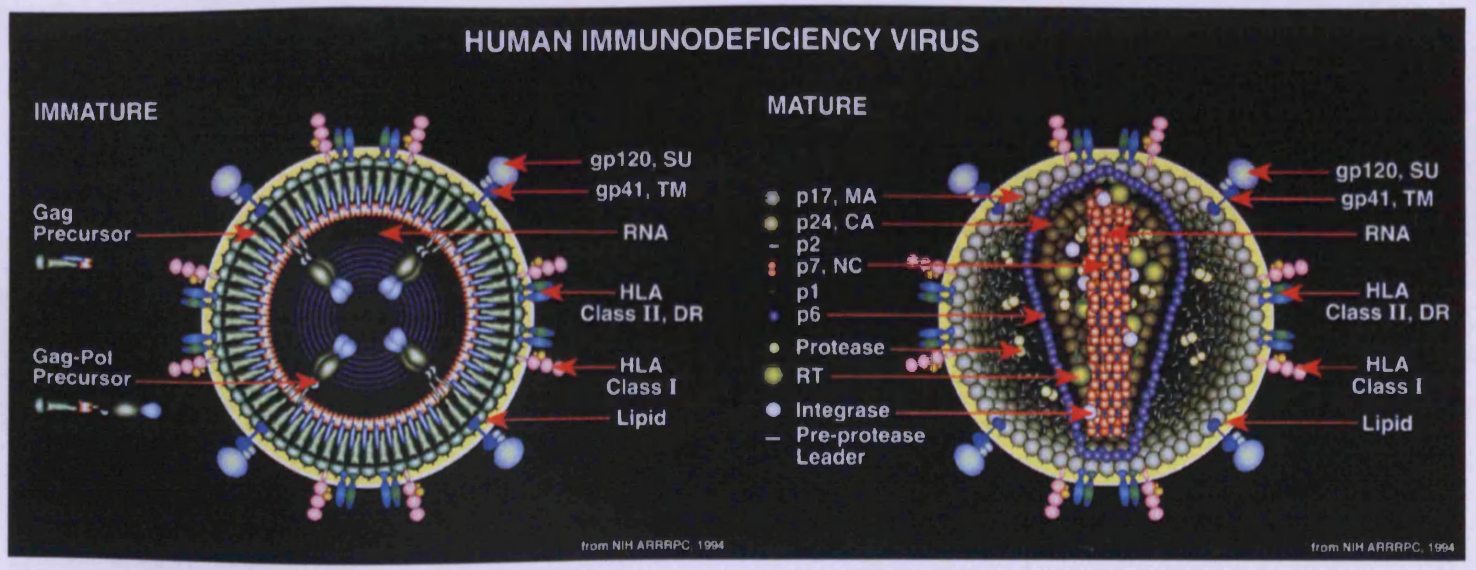


Figure 1.2. Morphology of *Retroviridae* virions

a. Morphology of betaretrovirus, lentivirus and gammaretrovirus particles. Betaretroviruses have an eccentrically placed round core (formerly known as type B morphology); lentiviruses have an elongated cone-shaped core and gammaretroviruses have a round core centred in the middle of the particle (formerly known as type C morphology). (Reproduced with permission from Principles of Virology 2nd edition, SJ Flint; LW Enquist; VR Racaniello; AM Skalka, ASM Press, Washington DC).

b. Morphology of HIV-1 virions before and after maturation. Immature HIV-1 virions resemble gammaretroviral particles; mature HIV-1 virions have an elongated cone-shaped core. (Image obtained from the International Committee on Taxonomy of Viruses database (ICTVdb) with kind permission from Dr Cornelia Buechen-Osmond. This image was modified from a model by Drs Louis Henderson and Larry Arthur obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH.)

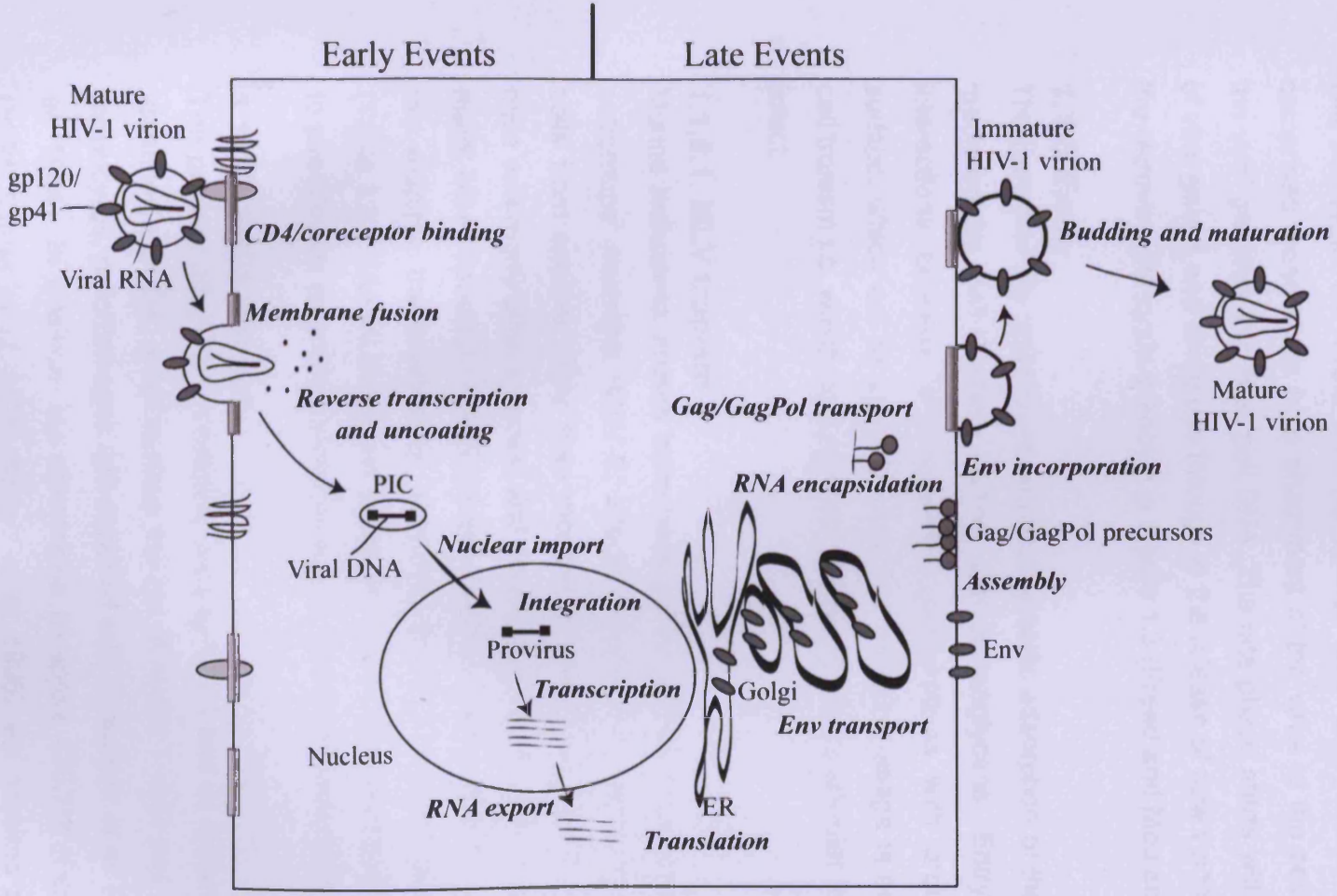


Figure 1.3. The retrovirus life cycle

The retrovirus life cycle is described in detail in section 1.1. of the introduction. An overview of the early and late events of the retrovirus life cycle is shown here for HIV-1. (Reproduced from Freed *et al*, Retrovirology **3**: 77, 2006, with kind permission from the publisher BioMed Central Ltd).

1.1.4. The retroviral life cycle

The retroviral life cycle can be divided into two phases: the early phase comprises those steps from attachment of the virus to the cell, to integration of the viral genome into host cell DNA. The late phase starts with the transcription of viral genes and continues through to the release of new virions. An overview of the retroviral life cycle is shown in Figure 1.3 (Freed and Mouland 2006).

1.1.5. Entry

The retroviral life cycle starts with non-specific adsorption of the virus to the cell, mediated by cell surface lectins and proteoglycans. Entry occurs through interactions between viral envelope glycoproteins with proteins on the cell surface, which act as specific receptors. Receptor-usage is one determinant of cell tropism i.e. which host species and cell types a particular retrovirus is able to infect.

1.1.5.1. MLV tropism

Murine leukaemia viruses have been subclassified into four different host range subgroups: ecotropic (MLV-E, infects mouse cells); xenotropic (MLV-X, infects cells from species other than mouse), amphotropic (MLV-A, infects cells from mice and many other species) and polytropic (MLV-P, infects cells from mice and many other species). The receptors used by these different subgroups of MLV are multiple transmembrane spanning proteins that function as transporters (Table 1.2.). The MLV-A envelope glycoprotein is frequently used in gene therapy to pseudotype retroviral cores, because it allows efficient entry into human cells.

1.1.5.2. HIV-1 tropism

The primary attachment receptor used by HIV-1 and its close relatives has been identified as CD4, a cell surface marker of helper T cells and also expressed on monocytes, macrophages and dendritic cells (Dalgleish *et al.* 1984; Klatzmann *et al.* 1984). In addition, the chemokine receptors CXCR4 (Feng *et al.* 1996) or CCR5 (Deng *et al.* 1996; Dragic *et al.* 1996) are required as co-receptors to mediate membrane fusion.

1.1.5.3. pH-dependent versus pH-independent fusion

HIV-1 and MLV-A enter cells by fusion with the plasma membrane, which is pH-independent (McClure *et al.* 1988). Whereas the surface unit (SU) of the viral envelope glycoprotein is responsible for recognising a specific receptor, fusion is mediated by the transmembrane (TM) domain. It occurs as a result of conformational changes in the envelope glycoprotein following receptor binding.

MLV-E enters cells by fusion with endosomal membranes (McClure *et al.* 1990). This process is pH-dependent, as low pH drives the conformational changes in envelope proteins needed for fusion. Vesicular Stomatis Virus (VSV), whose G protein is frequently used to pseudotype retroviral particles also gains entry by endocytosis (Fan and Sefton 1978).

1.1.6. Uncoating

Following membrane fusion, the retroviral capsid is released into the cytoplasm. Subsequent early post-entry events are considered the least well-understood of the retroviral life cycle. Uncoating, the (partial) disintegration of the retroviral capsid, is necessary for the generation of a reverse transcription complex (RTC). Capsid remains tightly associated with MLV RTC (Fassati and Goff 1999), whereas HIV-1 sheds its capsid shell rapidly after entry (Fassati and Goff 2001). Other viral proteins associated with the two copies of ssRNA genome in the RTC include RT and IN (plus NC, MA and Vpr in the case of HIV-1). Several cellular proteins have also been shown to be associated with MLV and HIV-1 RTCs.

Table 1.1. Genera of the *Retroviridae* family

Genus	Typical species	Genome
1. Alpharetrovirus	Avian Leukosis Virus (ALV)	Simple
2. Betaretrovirus	Mouse Mammary Tumour Virus (MMTV)	Simple
3. Gammaretrovirus	Murine Leukaemia Virus (MLV)	Simple
4. Deltaretrovirus	Bovine Leukaemia Virus	Complex
5. Epsilonretrovirus	Walleye Dermal Sarcoma Virus	Simple
6. Lentivirus	Human Immunodeficiency Virus Type I (HIV-1)	Complex
7. Spumavirus	Simian Spumavirus	Complex

Table 1.2. Tropism of MLV subgroups and receptor usage

Subgroup	Receptor	Examples
Ecotropic	Cationic amino acid transporter 1 (CAT-1) (Kim *et al.* 1991)	Moloney MLV, Friend MLV, AKR MLV (AKV)
Xenotropic	Rmc1 (XPR) (Battini *et al.* 1999; Tailor *et al.* 1999; Yang *et al.* 1999)	Endogenous MLVs
Amphotropic	Inorganic phosphate transporter 2 (Pit-2) (van Zeijl *et al.* 1994)	Exogenous MLVs
Polytropic	Rmc1 (XPR) (Battini *et al.* 1999; Tailor *et al.* 1999; Yang *et al.* 1999)	Endogenous MLVs Mink Cell Focus-inducing (MCF) viruses

1.1.7. Reverse transcription

Reverse transcriptase (RT) is a unique enzyme; it not only has RNA- and DNA-dependent DNA polymerase activity (Baltimore 1970; Temin and Mizutani 1970), but also possesses an RNase H domain that degrades genomic RNA after it has been copied into DNA. An overview of reverse transcription is shown in Figure 1.4 (Katz and Skalka 1994).

It starts near the 5'end of the retroviral genome. Minus strand synthesis is initiated from a tRNA primer bound to the primer binding site (PBS), located just downstream of U5. Minus strand (-) strong stop DNA approximately 100bp in length is generated, followed by RNase H digestion of the RNA portion of the RNA:DNA duplex. Complementary hybridisation of (-) strong stop R sequence to the 3' R sequence on the same or on the other RNA strand then occurs (1^{st} template exchange). Extension of minus strand synthesis follows. RNaseH digestion removes most plus strand RNA, except at the polypurine tract (PPT) which is highly resistant to RNaseH digestion. PPT RNA serves as a primer for plus strand (+) strong stop DNA synthesis. HIV also contains a central PPT (cPPT) for priming plus strand synthesis. The tRNA primer is removed. (+) strong stop DNA then hybridises to the complementary PBS on the minus strand (2^{nd} template exchange), leading to the completion of reverse transcription and generation of long terminal repeats (LTRs) at both ends of the dsDNA genome. The LTRs are essential for the subsequent integration reaction (Panganiban and Temin 1983) and for transcription of viral genes thereafter.

The template exchanges during reverse transcription are thought to be responsible for the high rate of recombination seen in retroviruses. Also, the RT enzyme has no proof reading activity. Whereas cellular DNA polymerase has an error rate of 1 misincorporation per 10^7 to 10^{11} nucleotides polymerised, the error rate of RT is thought to be as high as 1 in 10^4 to 1 in 10^6. The low fidelity of RT is responsible for the high in vivo mutation rate of retroviruses (Reviewed in: (Williams and Loeb 1992)).

The cellular protein APOBEC3G can interfere with HIV-1 reverse transcription. One mechanism by which it can do this is by inducing cytidine deamination in the

minus strand, resulting in hypermutation of reverse transcribed DNA (Reviewed in: (Holmes *et al.* 2007)). APOBEC3G is targeted for degradation by the HIV-1 encoded accessory protein Vif (Sheehy *et al.* 2003).

1.1.8. Nuclear import

Following the completion of reverse transcription the RTC becomes known as the pre-integration complex (PIC). Retroviral RTC/PIC are thought to be transported through the cytoplasm via the actin and microtubule cytoskeleton; passive diffusion cannot be responsible for their intracellular movement (Suzuki and Craigie 2007). The manner by which a PIC enters the nucleus has not as yet been completely resolved. Gammaretroviruses such as MLV can only infect dividing cells; breakdown of the nuclear membrane during mitosis therefore seems to be required for MLV PIC to access host chromatin. Lentiviruses such as HIV infect dividing and non-dividing cells equally well (Lewis *et al.* 1992). Import of HIV PICs into the nucleus most likely involves traversing the nuclear pore complex (NPC), although Vpr-mediated disruption of the nuclear envelope has also been proposed (de Noronha *et al.* 2001). All proteins in the cell that are destined for the nucleus contain nuclear localisation signals (NLS). The HIV PIC contain proteins such as MA, IN and Vpr in which NLS have been identified. The cPPT in the HIV genome has also been shown to facilitate nuclear entry (Zennou *et al.* 2000). However, the importance of these NLS has been called into question since it was shown that mutation or complete deletion of the four known NLS did not abolish HIV-1's ability to infect non-dividing cells (Dvorin *et al.* 2002; Yamashita and Emerman 2005). Capsid is now thought to be the main viral determinant of retroviral infectivity of non dividing cells (Yamashita *et al.* 2007). Cellular factors involved in the nuclear import of HIV-1 include importin 7 (Fassati *et al.* 2003) and a 3' CCA deficient tRNA species that is incorporated into HIV-1 virions (Zaitseva *et al.* 2006).

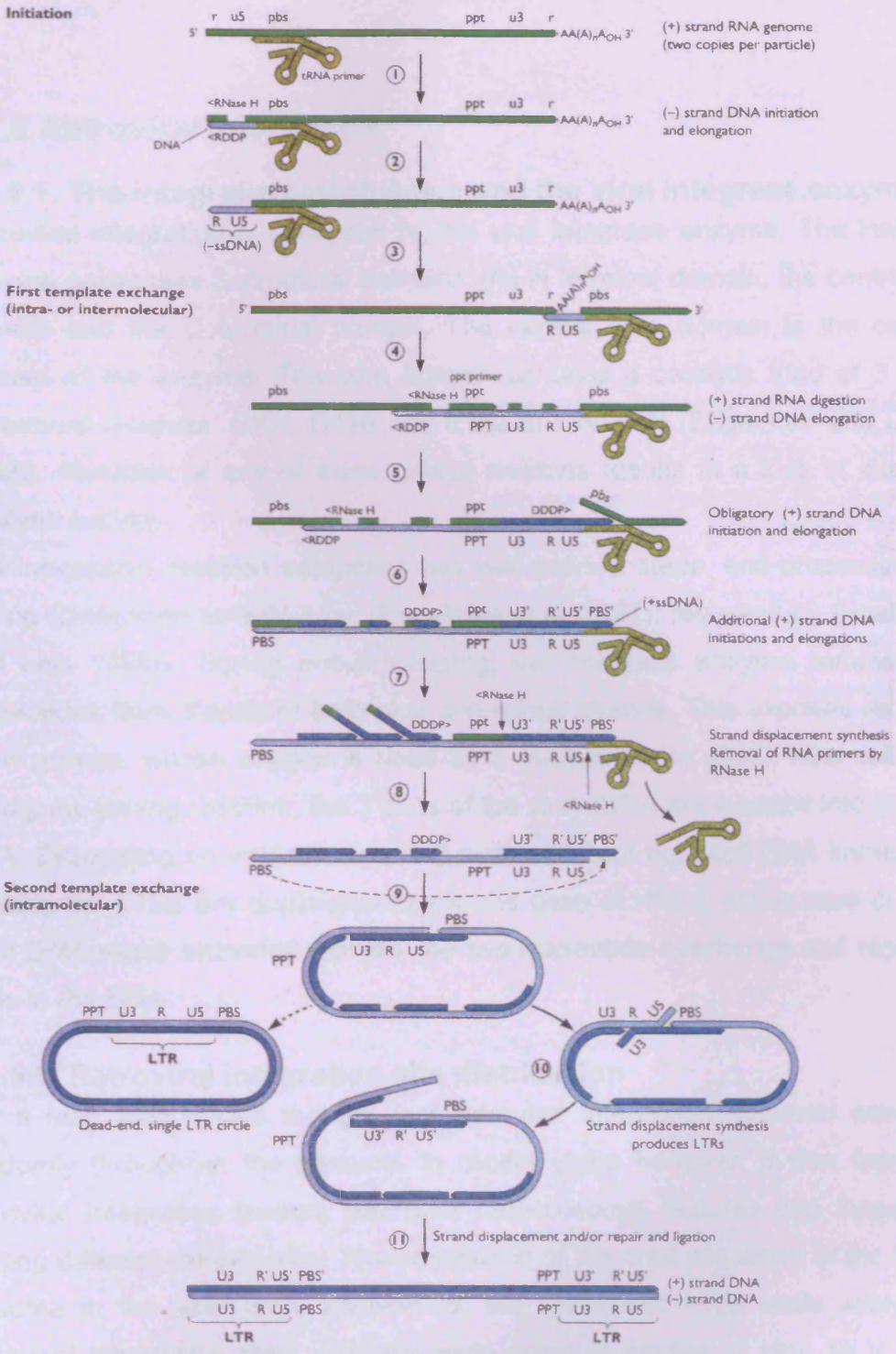


Figure 1.4. Overview of the reverse transcription process
RDDP is RNA-dependent DNA polymerisation; DDDP is DNA-dependent DNA polymerisation.
Reproduced from Katz *et al*, Annual Review of Biochemistry **63**: 133-173, 1994
Reprinted, with permission, from the Annual Review of Biochemistry, Volume 63 ©1994 by Annual Reviews www.annualreviews.org. Image obtained from a reproduction in Principles of Virology 2nd edition, SJ Flint; LW Enquist; VR Racaniello; AM Skalka, ASM Press, Washington DC.

1.1.9. Retroviral integration

1.1.9.1. The integration mechanism and the viral integrase enzyme

Retroviral integration is catalysed by the viral integrase enzyme. The integrase enzyme possesses 3 structural domains: the N terminal domain, the central core domain and the C terminal domain. The central core domain is the catalytic domain of the enzyme. The core domain contains a catalytic triad of 3 highly conserved residues. (D64, D116 and E152 in HIV-1 IN (Engelman and Craigie 1992)). Mutation of any of these critical residues results in a loss of integrase catalytic activity.

The integration reaction comprises two well-defined steps, end-processing and joining (Described for HIV-1 by: (Engelman *et al.* 1991); reviewed in: (Hindmarsh and Leis 1999)). During end-processing, the integrase enzyme removes two nucleotides from 3'ends of both plus and minus strands. This exposes recessed 3'OH groups, whose oxygen is used as a nucleophile to attack host cell DNA. During the joining reaction, the 3'ends of the viral cDNA are inserted into host cell DNA. Depending on viral species, 4-6 nucleotides of host cell DNA immediately flanking the LTRs are duplicated (5nt in the case of HIV-1; 4nt in case of MLV). Host DNA repair enzymes remove the two nucleotide overhangs and repair the gaps in the DNA.

1.1.9.2. Retroviral integration site distribution

For a long time it was thought that retroviral integration occurred essentially randomly throughout the genome. In recent years however, it was found that retroviral integration favours particular chromosomal features and these differ among different retroviruses. The completion of the draft sequence of the human genome in the year 2000 allowed for the systematic large scale analysis of retroviral integration sites. Genome wide mapping studies of HIV, MLV, ASLV and foamy virus integration sites have been conducted. Below follows a summary of the distinct target sites favoured by these different retroviruses.

1.1.9.2.1. Lentiviruses (HIV-1 and SIV)

The first genome wide mapping study of HIV-1 integration sites in the SupT1 T cell line revealed that genes are favoured targets of HIV-1 integration (Schroder *et al.* 2002). 69% of HIV-1 integration sites were found to be in transcription units (TU), compared to 35% of control in vitro integration sites into naked DNA. Integration site selection was also influenced by transcriptional activity. Microarray analysis showed that the median expression levels of genes hosting HIV-1 integration sites were consistently higher than the median expression levels of all genes on the array. Genes transcribed as a result of HIV-1 infection were particularly favoured targets for HIV-1 integration.

These preferences were confirmed by a study that analysed HIV-1 vector integration sites in human primary cells (Mitchell *et al.* 2004). Simian immunodeficiency (SIV) vector integration sites were analysed in the progeny of transduced rhesus macaque haematopoietic stem cells (Hematti *et al.* 2004). SIV like HIV favoured integration into TU.

1.1.9.2.2. Murine Leukaemia Virus (MLV)

20.2% of MLV integrations occurred 5kb upstream or downstream of transcription start sites compared to 4.3% of random integration sites (Wu *et al.* 2003). CpG islands are regions abundant in the CpG dinucleotide, which are undermethylated and associated with gene regulatory regions. MLV strongly favours integration near CpG islands, but HIV does not. Microarray analysis revealed that MLV preferentially integrates near transcription start sites of actively transcribed genes. MLV, unlike HIV or ASLV, also has a strong preference for integration near DNase I hypersensitive sites (Lewinski *et al.* 2006). These are nucleosome depleted chromosomal regions associated with gene regulatory elements. They are enriched near promoters, the 5'ends of TU and CpG islands.

1.1.9.3. Integration target site selection

Several models have been proposed to account for the distinct integration target site preferences of different retroviruses (Ciuffi and Bushman 2006).

The chromosome accessibility model proposes that open chromatin is a favoured target for PICs because it is most accessible. This model might account for the fact that centromeric heterochromatin is a consistently unfavourable environment for retroviral integration. DNAse I hypersensitive sites are markers for accessible chromatin, yet these sites are only favoured by MLV. It was thought that maybe cell cycle effects could explain the differences in integration preferences between HIV and MLV. However, no difference in HIV integration site selection was seen between dividing and cell cycle arrested cells (Ciuffi *et al.* 2006). The tethering model proposes that the PICs of different retroviruses interact with different cellular factors that guide them to specific sites on chromosomal DNA. This could account for integration preference in the proximity of transcription start sites (MLV), within genes (HIV) and in transcriptionally active regions (HIV and MLV). Although host cell DNA primary sequence seems of little importance in target site selection, distinct base preferences at the sites of integration and proximal sequences were observed in one study (Holman and Coffin 2005). This would suggest that integrase does recognise a sequence pattern upon integration.

1.1.9.4. Viral and cellular factors involved in retroviral integration and their possible role in target site selection

1.1.9.4.1. Viral determinants

Not surprisingly, integrase as the catalyst of the retroviral integration process is the main viral determinant of target site selection. A recent study by the Bushman lab demonstrated this very convincingly (Lewinski *et al.* 2006). Emerman's group had previously generated a set of chimeric HIV viruses, one in which HIV IN was replaced by MLV IN (HIVmIN); another one in which HIV Gag and IN were replaced by their MLV counterparts (HIVmGagmIN); and a third one in which only HIV Gag was replaced by MLV Gag (HIVmGag)(Yamashita and Emerman 2004; Yamashita and Emerman 2005). Only HIVmIN could infect non-dividing cells, meaning MLV Gag imposes the requirement for cell division on the HIVmGag and HIVmGagmIN chimeras. The integration site preferences of the different chimeric viruses were compared. The HIVmGagmIN pattern most closely resembled that of MLV, followed by that of HIVmIN; the integration site

preference of HIVmGag differed from that of both MLV and HIV. In conclusion, integrase is thus the main viral determinant of integration specificity, but Gag derived proteins also have some influence.

1.1.9.4.2. Cellular determinants

Several cellular factors such as HMGA1(high-mobility-group protein A1), BAF (Barrier to Autointegration Factor), EED, p300, Ini-1 (Integrase Interactor 1) and LEDGF/p75 are known to be associated with PICs and these in turn could influence retroviral integration and target site selection (Ciuffi and Bushman 2006). Recently, it was reported that HIV-1 integration is dependent on the integral inner-nuclear-membrane protein emerin and its LEM (LAP, emerin, MAN) binding partner BAF (Jacque and Stevenson 2006). Emerin and BAF are required for the association of HIV-1 PICs with chromatin. When emerin or BAF are knocked down by RNA interference, HIV-1 PICs are sequestered in the nuclear matrix where they are present as episomal 1'LTR and 2'LTR circles. Infection by MLV, which is restricted to dividing cells, is dependent on Lap2α and BAF (Suzuki *et al.* 2004), but not emerin (Jacque and Stevenson 2006). MLV and HIV thus use different inner nuclear membrane proteins to co-localise with chromatin. Lap2α is only required for integration by HIV-1 carrying wild-type envelope proteins, but not when it is pseudotyped with VSV-G. The route of entry into the cell thus also seems to play a role in the requirement for different nuclear proteins mediating integration. Lap2α & BAF and emerin & BAF are required for chromatin association of MLV PICs and HIV PICs, respectively, but they are not known to influence integration target site selection by the two different viruses. In this study it was also found that HMGA1 was not required for MLV or HIV integration.

So far, the only cellular factor known to influence integration site selection is lens epithelium-derived growth factor LEDGF/p75 (Ciuffi *et al.* 2005). LEDGF/p75 is specific for lentiviral targeting. It strongly binds to lentiviral integrase as well as chromosomal DNA. Ciuffi et al knocked down LEDGF/p75 expression in 3 different cell types (293T, Jurkat, HOS) and investigated whether this altered HIV-1 integration site preferences in these cells (Ciuffi *et al.* 2005). HIV

integration was still favoured in transcription units in LEDGF/p75 depleted cells, but it was markedly reduced (56% of HIV integration sites were in TU in LEDGF/p75 depleted cells versus 67% in control cells). These data suggest that LEDGF/p75 does contribute to targeting HIV to transcription units but is not the sole factor in doing so.

1.1.10. Retroviral gene expression

In both simple and complex retroviruses gene expression is initiated from the 5' long terminal repeat (LTR). The LTRs are subdivided into three regions, namely the U3, R and U5 regions. Transcription is initiated from the U3/R boundary (designated at +1 in Figure 1.5.). Retroviral gene expression relies on host cell transcriptional machinery; the RNA polymerase II (RNAPII) complex assembles at the TATAA box, located 30bp upstream of the transcription start site in both the MLV and HIV-1 LTRs. The LTR U3 region in particular contains binding sites for many different cellular transcription factors that regulate activity of the viral promoter. Regulation of MLV and HIV-1 LTR activity shall be discussed in detail.

1.1.10.1. MLV LTR activity

The Moloney MLV (Mo-MLV) enhancer is located in the U3 region, 342 to 154bp upstream of the transcription start site. It consists of two 75bp direct repeats plus an additional upstream glucocorticoid response element (GRE) and downstream mammalian type-C retrovirus enhancer factor 1 (MCREF-1) binding site (Manley *et al.* 1993). The Mo-MLV enhancer not only controls transcriptional activity of its own promoter, it can also trans-activate nearby cellular promoters. The implications of this will be discussed in the section on MLV pathogenesis.

The Mo-MLV enhancer contains numerous binding sites for cellular transcription factors. Some of these overlap, so a particular enhancer sequence can bind multiple proteins. This property allows tissue specific regulation of Mo-MLV LTR activity. The two most important binding sites in the Mo-MLV direct repeats are the Core Binding Factor (CBF) site (Wang *et al.* 1993) and the two Ets-1 binding sites (LVb and LVc) (Gunther and Graves 1994) that flank it. Both these proteins are highly enriched in haematopoietic cells. The CBF, LVb and LVc sites are

highly conserved among Type C retroviruses (Golemis *et al.* 1990). Mutations in the CBF site reduces LTR transcriptional activity in T and B cells, but not in fibroblasts (Speck *et al.* 1990). Two NF-1 sites are also contained within the direct repeat (Reisman 1990), mutations of these sites affect transcription in fibroblasts but not in haematopoietic cells. Furthermore, one GRE, an E box motif which binds basic helix-loop-helix (HLH) transcription factors such as ALF-1 (Corneliussen *et al.* 1991; Nielsen *et al.* 1994) and an MCREF-1 (Manley *et al.* 1993) binding site are present in each Mo-MLV direct repeat.

1.1.10.2. HIV-1 LTR activity

The following reviews on HIV-1 transcription were consulted to write this section: (Fred C. Krebs 2001); (Stevens *et al.* 2006); (Williams and Greene 2007). HIV-1 LTR gene expression relies both on cellular transcription factors whose expression is enriched in T cells and monocytes/macrophages, as well as a virus encoded transactivator protein Tat. The contribution of both these factors shall be discussed in turn.

1.1.10.2.1. Cellular transcription factors

The HIV-1 U3 region can be subdivided into three functional regions: the core promoter; the basal enhancer and the modulatory region (Figure 1.5.b.). The core promoter contains the TATAA box and three Sp1 binding sites. Sp1 is a ubiquitously expressed cellular transcription factor; it is thought to direct components of RNAPII to the TATAA box. The HIV-1 basal enhancer contains an AP-2 (Perkins *et al.* 1994) and two NFκB binding sites (Nabel and Baltimore 1987). Nuclear factor κ B (NFκB) is the most important cellular transcription factor in HIV-1 gene expression; its role shall be discussed in detail below.

The modulatory region contains an extraordinary number of transcription factor binding sites, including those for Ets-1, GATA-3, LEF-1/TCF-1α; NF-IL6 (C/EBP), AP-1 and NFAT. The latter two are inducible T cell transcription factors; they were shown to synergise/cooperate with NFκB in activating HIV-1 gene expression (Kinoshita *et al.* 1997; Yang *et al.* 1999). The effect of most of these transcription factors has only been assessed by quantifying HIV-1 driven reporter

gene expression in transient systems and their role in HIV-1 viral gene expression in vivo is unknown.

In HIV-1, cellular transcription factors are required for basal transcription, Tat-dependent transcription and the reactivation of HIV-1 from latency. HIV-1 reproduces most efficiently in activated T cells, suggesting the transcription factors present in this cell subset is required for efficient HIV-1 gene expression.

In unstimulated T cells, NFκB (p65/p50) is sequestered in the cytoplasm by IκB. Upon T cell stimulation by e.g. T cell receptor ligation, cytokine signalling by IL-1β, IL-7 or TNF-α or mitogen stimulation with phorbol myristate acetate (PMA), IκB is phosphorylated and targeted for degradation, whereupon NFκB p65/p50 translocates to the nucleus. In unstimulated cells, NFκB p50 homodimers occupy the NFκB binding sites in the basal enhancer. Histone deacetylase 1 (HDAC-1) is bound to p50 homodimers; this enzyme maintains the histones nucleosome 1 (nuc-1) and nuc-0 in a deacetylated state (Zhong *et al.* 2002; Williams *et al.* 2006). These nucleosomes inhibit the binding of transcription factors and RNAPII components to the HIV-1 promoter. When the p65/p50 NFκB heterodimer enters the nucleus following T cell stimulation, it displaces the p50 homodimer from the NFκB core enhancer sites. p65/p50 binds the histone acetyl transferase (HAT) p300. p300 acetylates nuc-1. This leads to the recruitment of chromatin remodelling complexes that disrupt nuc-1, enabling the initiation of transcription (Verdin *et al.* 1993; Van Lint *et al.* 1996).

In the absence of Tat, transcription is successfully initiated by cellular transcription factors. However, elongation is very inefficient with 90% of transcripts terminating within 60bp of the transcription start site. NFκB not only directs transcriptional initiation, it also has a role in elongation in the absence of Tat (West *et al.* 2001). NFκB p65/p50 recruits TFIIH, which phosphorylates the C terminal domain (CTD) of RNAPII on serine 5 (Kim *et al.* 2006). NFκB directed elongation is crucial in the reactivation of HIV-1 from latency (Williams *et al.* 2007).

Figure 1.5.a. Transcription factor binding sites in the Moloney Murine Leukaemia Virus enhancer

The 594 bp Moloney Murine Leukaemia Virus long terminal repeat is depicted here. The enhancer is located in the U3 region, 342 to 154 bp upstream of the transcription start site at the U3/R boundary (indicated by + 1 and the arrow). It consists of two 75 bp direct repeats plus an upstream glucocorticoid response element (GRE) and downstream mammalian type-C retrovirus enhancer factor 1 (MCREF-1) binding site. Individual transcription factor binding sites in the direct repeat are shown in the enlargement. This diagram was adapted from Figure 7 in Manley *et al* J. Virol. 1993 **67**(4): 1967-1975. (Ets-1 is Ets binding site; CBF is core binding factor; bHLH is basic Helix Loop Helix; NF-1 nuclear factor 1). The TATA and CAAT elements are found at 30 and 81 bp upstream of the transcription start site, respectively.

Figure 1.5.b. Transcription factor binding sites in the Human Immunodeficiency Virus-1 (HIV-1) U3 region

The 633 bp HIV-1 LTR consists of U3, R and U5 regions. Transcription is initiated from the U3/R boundary (designated +1 and marked by the arrow). The Transactivation Response (TAR) element is encoded by the first 60 bp of R; it forms an RNA stem loop structure which binds Tat. The U3 region can be subdivided into three functional regions: the core promoter; the basal enhancer and the modulatory region. Some of the key transcription factor binding sites in the U3 region are shown in the enlargement. Additional transcription factor binding sites exist in R, U5 and the Gag leader sequence. The proviral genome is packaged into chromatin. The position of the two nucleosomes 0 and 1 on the HIV-1 LTR are shown. The acetylation status of these nucleosomes controls LTR gene expression.

COUP-TF is chicken ovalbumin upstream promoter transcription factor; RARE is retinoic acid response element; AP-1 is activator protein 1; Myb is Myb binding factor; NFAT is nuclear factor of activated T cells; GR is glucocorticoid receptor binding site; USF-1 is upstream stimulatory factor 1; Ets-1 is Ets-1 binding factor; LEF-1is Lymphoid enhancer-binding factor 1; C/EBP (NF IL-6) is CCAAT/enhancer binding protein; AP-2 is activator protein 2; NFκB is nuclear factor kappa B.

This diagram of the HIV-1 5'LTR was adapted from Figure 1. in Stevens *et al* Aids 2006 **12**(3) 253-259.

1.1.10.2.2. Tat-dependent transcription

As just discussed, some elongated transcript is transcribed from the HIV-1 LTR in the absence of Tat. The 101 amino acid Tat protein is encoded by a multiply spliced transcript of two exons. It contains a NLS, so is translocated to the nucleus after translation. Tat binds the TAR loop (Kao *et al.* 1987). This is an RNA element encoded by the first 60bp of the R region; it forms a highly stable secondary stem loop structure which is recognised by Tat. Tat recruits the positive-acting transcription elongation factor b (p-TEFb) to the HIV-1 promoter (Garcia-Martinez *et al.* 1997; Mancebo *et al.* 1997). The two most important components of p-TEFb are human cyclin T1 and CDK9 (Wei *et al.* 1998). CDK9 phosphorylates the CTD of RNAPII on serine 2 and 5 (Zhou *et al.* 2000). CTD hyperphosphorylation greatly increases processivity of RNAPII and allows efficient elongation of transcripts. Tat also recruits additional HATs to the promoter, further improving DNA access.

1.1.11. 5' and 3' end processing

All proviral transcripts are capped at their 5'end, again using cellular capping machinery. HIV-1 Tat interacts with mRNA capping proteins, which stabilises nascent RNA transcripts. The poly(A) signal in both MLV and HIV-1 is found in the R region of the LTR. Transcription proceeds past this site; transcripts are then processed by nucleases and polyadenylated at the R/U5 boundary. Polyadenylation signals are thus present in the both the 5' and 3' LTRs. In HIV, 5' poly(A) site usage is suppressed by U1 snRNP interaction with the major splice donor (MSD) (Ashe *et al.* 1997). In MLV, mutation of the downstream SD does not affect the activity of the 5' polyadenylation signal (Furger *et al.* 2001). It is thought that the inability of MLV to actively suppress 5' polyadenylation, necessitates the use of a weaker polyadenylation signal. In addition, upstream sequence elements in the U3 region enhance 3' polyadenylation; these are found just downstream of the TATAA box in HIV-1 and dispersed throughout the U3 region in MLV.

1.1.12. Splicing and nuclear export

Cellular mRNAs that contain introns or splice sites are normally retained in the nucleus. Nuclear export of unspliced viral RNA is however required for the expression of Gag-pol proteins and for packaging of full length genomic RNA into new virions.

RNA transcripts of simple retroviruses come as unspliced full length transcripts and singly spliced envelope encoding transcripts. Simple retroviruses do not encode proteins that mediate splicing or RNA nuclear export. The ratio of unspliced to spliced transcripts is however important so that at a later stage proportionate quantities of viral proteins are made. Simple retroviruses are thought to contain RNA elements that allow transport of full length RNA from nucleus to cytoplasm, e.g. the constitutive transport element (CTE) in the Mason-Pfizer Monkey betaretrovirus. It was recently reported that the Mo-MLV packaging signal is involved in nuclear export of full length RNA (Smagulova *et al.* 2005).

Transcripts of complex retroviruses such as HIV-1 can undergo multiple alternative splicing events. HIV-1 Tat, Rev and Nef are encoded by multiply spliced transcripts; these are exported from the nucleus to the cytoplasm. Therefore, these proteins are expressed early after HIV-1 infection of the cell.

Nef has many roles in HIV-1 infection (Steffens and Hope 2001). It downregulates CD4 and other cell surface proteins such as MHC Class I. This probably has a role in superinfection resistance and the recognition of infected cells by cytotoxic T cells. Nef also enhances virion infectivity. The Rev protein is critically important in the nuclear export of full length and incompletely spliced HIV-1 transcripts, which encode HIV-1 structural proteins (Malim *et al.* 1989). Rev recognises the Rev-responsive element (RRE) present in these transcripts. The RRE is an RNA element in the env coding region that folds into a complex secondary structure in a manner analogous to the TAR loop. Rev interacts with cellular transport proteins to mediate nuclear export of RRE containing transcripts.

1.1.13. Late events in the retroviral life cycle: Translation, assembly, budding and maturation

Retroviral virions are assembled from three polyprotein precursors, namely Gag, Gag-pol and Env. The Gag polyprotein precursor is translated from unspliced viral RNA; it contains MA, CA and NC proteins. Gag and pol genes are separated either by a stop codon, or they may be in a different reading frame. Gag-pol fusion proteins are produced by either read through (termination suppression i.e. the stop codon is read as a coding codon) in the case of MLV or by ribosomal frameshifting in HIV-1. In this way, 5 to 10% of Gag polyproteins synthesised are Gag-pol fusion proteins (Shehu-Xhilaga *et al.* 2001).

Assembly of gammaretroviruses and lentiviruses occurs at the plasma membrane. Though HIV-1 assembly in $CD4^+$ T cells occurs at this site (Palmer *et al.* 1985), controversy existed surrounding the site of HIV-1 assembly in macrophages. Some studies suggested this occurred in late endosomes in these cells (Pelchen-Matthews *et al.* 2003); more recently it was shown these were in fact invaginations of the plasma membrane (Deneka *et al.* 2007).

Membrane targeting (M) domains are contained within the matrix (MA) sequence of Gag. Gag and Gag-pol polyproteins localise to the plasma membrane following myristoylation of their N terminal glycine. The Env polyprotein is translated from a spliced transcript. The N terminus of Env contains a short hydrophobic signal peptide that docks it to the surface of the rough endoplasmic reticulum (RER) and mediates its translocation into the lumen of the RER. The C terminus of Env is anchored into the RER membrane. Env reaches the cell surface via the cell's secretory pathway. On its way there it undergoes a number of modifications: it is glycosylated; it assembles into oligomers and it is cleaved by cellular proteases into its two subunits SU and TM.

Gag polyprotein is the driving force of retroviral assembly (Gottlinger 2001). In most retroviruses, Gag alone is sufficient to mediate budding. Virus like particles (VLPs) are efficiently generated in the absence of Env, but Env incorporation into virions is essential for them to be infectious. Gag polyprotein oligomerisation occurs through interaction (I) domains in NC. The NC portion of Gag also

contains an RNA-binding domain that recognises the packaging signal and ensures genomic RNA packaging into virions. Gag polyproteins are spatially orientated during assembly. MA (at the N terminus) and NC (towards the C terminus) make up the outermost and innermost layers of the retroviral protein shell, respectively.

Late (L) domains in Gag mediate retroviral budding (Reviewed in (Martin-Serrano 2007)). The key protein involved in this process is ESCRT-III. ESCRT stands for Endosomal Sorting Complex Required for Transport. These complexes normally function in the sorting of ubiquitinated cell membrane proteins for degradation in multivesicular bodies. MLV and HIV-1 recruit ESCRT-III in different manners:

The MLV p12 domain of Gag, located between MA and CA, contains a PPXY motif. It becomes ubiquitinated and recruits an HECT ubiquitin ligase. HIV-1 p6 contains a PTAPP domain which interacts with Tsg101, a subunit of the ESCRT-I complex. HECT ubiquitin ligase and ESCRT-I are thought to function as adaptor proteins that allow the L-domains of MLV and HIV-1, respectively, to interact with ESCRT-III, enabling retroviral budding. The virions that bud from the cell are immature and non infectious particles. Maturation to generate infectious virions occurs soon after budding and is dependent on viral protease (PR). PR cleaves Gag and Gag-pol polyproteins and gives rise to mature virions with distinct morphology. It was recently demonstrated that HIV-1 Vpu is required for the final release of mature HIV-1 virions from the plasma membrane. It counteracts a cellular factor that tethers these viral particles to the cell surface (Neil *et al.* 2008).

1.2. Retroviral pathogenesis

This thesis examines whether vectors derived from MLV and HIV-1 can transform cells and the significance of this for the safety of gene therapy. This section shall therefore focus on the pathogenesis of MLV and HIV in mice and humans, respectively. There is a long standing association between retrovirus infection and cancer. However, retroviruses cause a number of other diseases.

The most well known today is HIV induced immunodeficiency in humans. Immunodeficiency is also seen in a range of other species including mice (Du5H), cats (Feline Immunodeficiency Virus (FIV) and Feline Leukaemia Virus (FeLV)) and monkeys (SIV). Retrovirus associated neurological disorders have been described. HIV-1 and Human T lymphotropic virus 1 (HTLV-1) can cause AIDS dementia and tropical spastic paraparesis, respectively. FeLV-C and Equine Infectious Anaemia Virus (EIAV) can cause anaemia in cats and horses, respectively. This list is by no means exhaustive.

1.2.1. Acutely transforming retroviruses

At the start of the 20th century, it was shown that leukaemia and sarcoma could be induced in chickens within days of infection with transmissible agents. (Ellerman and Bang 1908; Rous 1979). These were later shown to be the retroviruses Avian Leukosis Virus (ALV) and Rous Sarcoma Virus (RSV), respectively. More such acutely transforming retroviruses were discovered over the years. The use of temperature sensitive mutants of RSV demonstrated that the virus was required for both initiation and maintenance of the transformed state (Martin 1970). The genome of RSV was shown to contain RNA sequences that are not present in replication competent but transformation defective avian leucosis viruses (Duesberg and Vogt 1970). This "extra" RNA carries a viral oncogene (v-onc). In 1976, the v-onc of RSV was shown to have originated from a cellular oncogene (c-onc) (Stehelin *et al.* 1976). This was a milestone finding in cancer research and led to the discovery of many more oncogenes. Acutely transforming retroviruses at some point captured c-onc sequences by

recombination. They transduce oncogenes into the cells they infect and can rapidly transform them in this way.

1.2.2. Chronically transforming retroviruses with special emphasis on Murine Leukaemia Viruses

Chronically transforming retroviruses do not transduce oncogenes. They too were found to induce cancer in susceptible hosts, albeit with a longer latency period. About a dozen different murine leukaemia virus (MLV) strains have been isolated over the years. Most cause haematopoietic tumours in mice; the type of tumour depends on the isolate (Reviewed in (Schiff and Oliff 1986)). Endogenous AKV, found in AKR inbred mice, was the first isolated strain of MLV. Serial passage of AKV through newborn mice resulted in the Gross Passage A strain of MLV. AKR mice succumb to T cell leukaemia/lymphoma within 10-12 weeks of birth; neonatal Gross A infection has the same effect. Neonatal Moloney MLV (Mo-MLV) infection causes T cell lymphoma. Mo-MLV infection causes acute pro-monocytic leukaemia (APML) in adult BALB/c mice undergoing a chronic inflammatory reaction in response to pristane treatment (Shen-Ong and Wolff 1987). Friend MLV (Fr-MLV), the replication competent component of Friend virus, causes mostly erythroleukaemias. The viral determinants responsible for disease specificity by these different strains of MLV shall be discussed.

1.2.2.1. Mo-MLV leukaemogenesis

Neonatal Mo-MLV infection induces overt T cell lymphoma with a latency period of 3-4 months in C57BL and BALB/c mice. However, as soon as 4-6 weeks post-infection pre-leukaemic abnormalities such as defects in bone marrow haematopoiesis, splenomegaly and thymic atrophy can already be observed (Davis *et al.* 1987; Li and Fan 1990; Bonzon and Fan 1999). These pre-leukaemic changes have been extensively studied but it is not clear what their role in leukaemogenesis is. It is thought these events are a result of chronic viraemia in infected mice.

1.2.2.2. Generation of Mo-Mink Cell Focus inducing (Mo-MCF) viruses and their role in leukaemogenesis

Mo-MLV infection in mice is accompanied by the generation of Mo-MCF viruses (Hartley *et al.* 1977). Mo-MCF viruses result from env recombination between sequences of exogenous ecotropic MLV and endogenous polytropic MLV; the latter are present in the mouse genome. Mo-MCF viruses have a polytropic host range. Normally, cells that have already been infected by ecotropic MLV cannot become re-infected by this virus, a phenomenon known as superinfection resistance, usually caused by viral envelope proteins occupying the receptor. Polytropic Mo-MCF viruses can superinfect Mo-MLV infected cells because they use a different receptor for entry. There is weak interference with superinfection by polytropic MLV (Marin *et al.* 1999), so Mo-MCF viruses can superinfect Mo-MCF-virus infected cells too. This can result in the accumulation of numerous provirus copies in a single cell. Logically, clones will emerge in which several cooperating oncogenes have been activated by insertional mutagenesis. This is consistent with the idea that several oncogenes, regulating distinct pathways, are required for cell transformation (Reviewed in (Hahn and Weinberg 2002)).This would explain why MLV variants that generate MCF viruses, induce lymphoma more efficiently (Brightman *et al.* 1991). Mo-MLV leukaemogenesis is also more efficient in mice compared to rats. The rat genome does not encode endogenous polytropic MLV, preventing MCF virus generation. However, Mo-MLV infected rats do develop lymphoma, demonstrating that MCF viruses are not strictly required for this.

1.2.2.3. Viral determinants of pathogenesis outside the LTR

As discussed in section 1.2.2.4., retroviral enhancers are the main determinant of MLV pathogenicity. However, determinants elsewhere in the virus exist. Mo-MLV, but not Fr-MLV, induces promonocytic leukaemia in adult pristane-treated mice. Unlike the differences in disease-type seen after neonatal infection with these 2 viruses, the U3 region was shown not to be responsible for this (Wolff and Koller 1990). Analysis of reciprocal chimeras of these viruses demonstrated that

determinants in the Ψ-gag-PR and env regions of Mo-MLV were responsible for leukaemia induction in this model instead (Mukhopadhyaya *et al.* 1994).

1.2.2.4. Proviral insertional mutagenesis of cellular proto-oncogenes: Retroviral enhancers as determinants of disease-type and latency

As described in the introduction to this section on MLV pathogenesis, different strains cause different types of haematological malignancies. The first clue as to what determines disease-specificity was given by the generation of chimeric viruses between Mo-MLV which causes T cell lymphoma and Fr-MLV which causes erythroleukaemia (Chatis *et al.* 1983). The LTRs of Mo-MLV were replaced by those of Fr-MLV, and this virus caused predominantly erythroleukaemias. The reverse chimera caused mainly T cell lymphoma. Later experiments showed that replacing just the enhancer elements (direct repeats) was sufficient to change disease-specificity (Li *et al.* 1987). In addition, the greater the number of direct repeats present in the enhancer, the shorter the latency period. This is consistent with the idea that enhancer activity is controlled by the presence of cell specific transcription factors. By implication, erythroid specific transcription factors bind the Friend enhancer, whereas T cell specific transcription factors bind the Moloney enhancer. The importance of individual transcription factor binding sites in the direct repeat was investigated next (Speck *et al.* 1990). The sites shown in Figure 1.5.a. were mutated one by one and their effect on Mo-MLV pathogenicity was tested. Mutations in at least 2 out of 3 GREs, in both LVb (Ets-1) sites, in 2 out of 4 NF-1 sites or in both CBF sites delayed the latency period of disease induction. In addition, mutations in the CBF/LVb sites changed the disease-specificity in favour of erythroleukaemia. CBF is highly expressed in T cells; a later study showed that the higher the affinity of CBF for its binding site, the more rapid is disease induction (Lewis *et al.* 1999). Mutations in both LVa, LVc or the promoter proximal GRE do not change disease specificity or latency. Deletion of the MCREF-1 site downstream of the direct repeats also relaxes disease specificity to erythroid and myeloid leukaemias (Manley *et al.* 1993).

These studies have important implications for viral vectors: The choice of LTR will determine transgene expression levels, which will vary between cell types. In terms of safety, cell types that express transcription factors whose binding sites are present in the enhancer will be at greater risk of transformation by insertional mutagenesis.

1.2.2.5. Mechanisms of insertional mutagenesis

Several ways in which chronically transforming retroviruses such as MLV can activate expression of cellular genes near integration sites have been described and are shown in Figure 1.6. If the cellular genes concerned are proto-oncogenes (c-onc), expression of which confers a survival and/or growth advantage onto the cell, this may be the initiating event in neoplastic transformation of this cell. In all mechanisms of insertional activation described below, the retroviral LTR is invariably involved. These cis-acting elements of the virus are retained in retroviral vectors, hence they have the potential to affect proto-oncogene expression in similar ways.

1.2.2.5.1. Promoter insertion

Promoter insertion (Figure 1.6.a) requires the provirus to integrate in the same transcriptional orientation as the gene. It was the first described mechanism whereby chronically transforming retroviruses could activate gene expression. In ALV-induced lymphomas in chickens, c-myc was activated by 3'LTR promoter insertion (Hayward *et al.* 1981). It was found that the 5'LTR of the provirus was deleted in most tumours. Usually the 3'LTR is inactive when the 5'LTR is transcriptionally active, because of promoter interference (Cullen *et al.* 1984). 5'LTR promoter insertion is also possible and has been described for c-myb activation by Mo-MLV (Shen-Ong *et al.* 1986). It involves generation of a fusion transcript, enabled by read-through through the 3'LTR polyadenylation site (Figure 1.6.c). This transcript is then spliced from a splice donor (SD) present in the virus to a splice acceptor (SA) in the downstream cellular gene. As long as provirus insertion occurs upstream of the gene or the first coding exon, a wild type protein will be produced.

1.2.2.5.2. Enhancer activation

Retroviral enhancers are able to activate cellular promoters over large distances. Evi1 activation by retroviral integration 90kb upstream was reported in a murine myeloid tumour (Bartholomew and Ihle 1991). C-myc activation by Mo-MLV integration 270kb downstream was demonstrated in a rat T cell lymphoma (Lazo *et al.* 1990). Enhancers can activate cellular promoters when the provirus has integrated both up- or downstream of the gene (Figure 1.6.b). Insertions can occur in any orientation with respect to gene transcription; however, in most cases upstream insertions occur in opposite transcriptional orientation, whereas downstream insertions occur in the same transcriptional orientation.

1.2.2.5.3. Disruption of mRNA destabilising elements

Figure 1.6.d shows proviral insertion into the 3'UTR of a gene. The 3'UTR of certain genes contains AU-rich elements (ARE) that destabilise mRNA. Proviral insertion in this region can result in premature termination of the transcript by the use of either a sense viral polyadenylation signal in the 5'LTR or a cryptic antisense polyadenylation signal in the viral enhancer (Uren *et al.* 2005). mRNA is stabilised and its half-life in the cytoplasm prolonged, resulting in increased protein production. This mechanism of insertional activation has been described for N-myc (van Lohuizen *et al.* 1989) and Pim-1 (Selten *et al.* 1985).

1.2.2.5.4. Inappropriate transcriptional termination

Proviral integration within a transcription unit can result in coding sequence disruption or premature termination of the transcript. This can in turn lead to production of a truncated, inactive or unstable protein. Important negative regulatory domains may be missing, generating constitutively active oncoproteins. For example, MoMLV insertions in the last intron of the Tpl-2 gene in rat T cell lymphomas results in the expression of an oncogenic carboxy terminal truncated Tpl-2 kinase protein (Ceci *et al.* 1997).

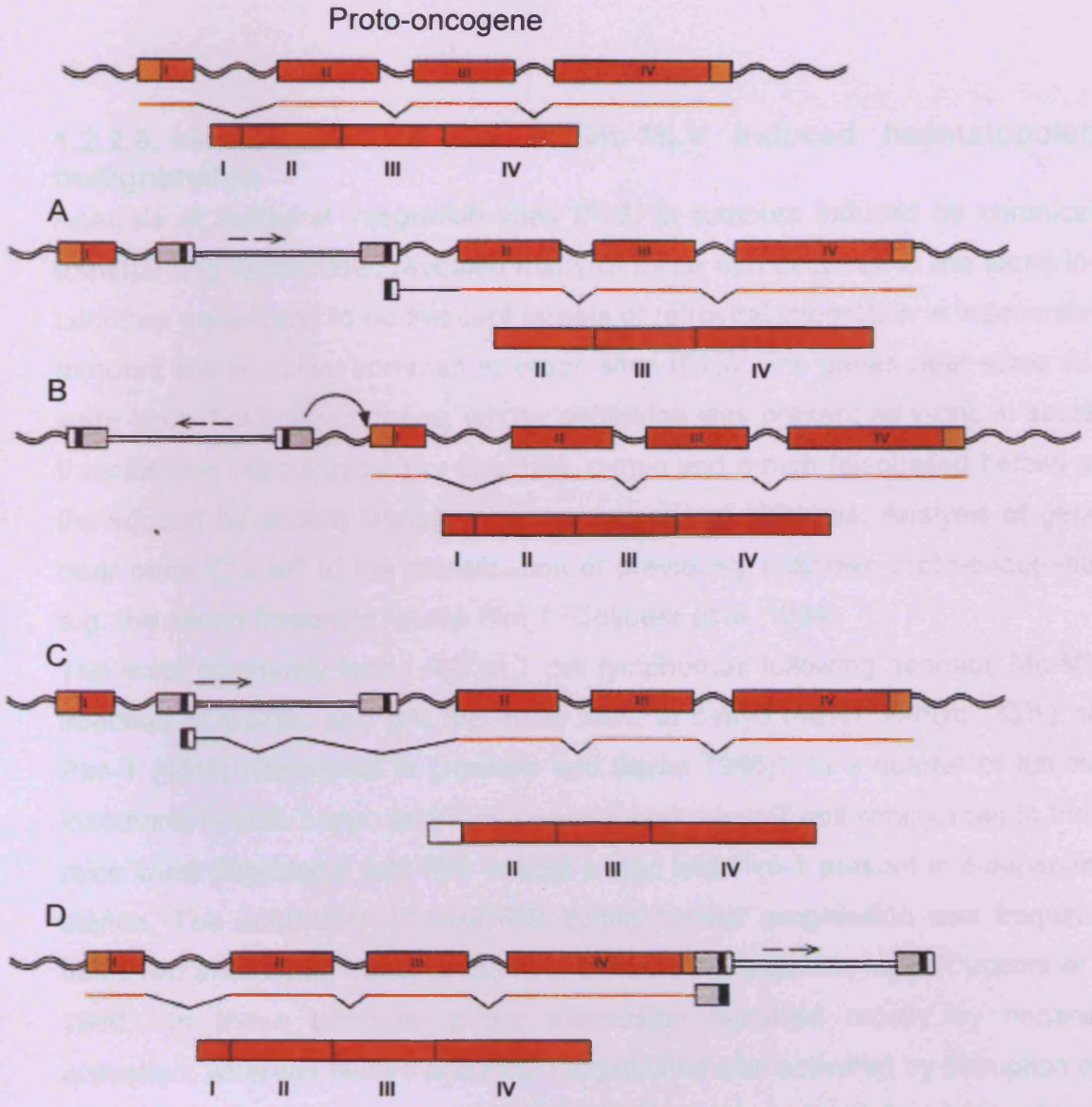


Figure 1.6. Mechanisms of insertional mutagenesis

A. Promoter insertion – provirus insertion has occurred in the same orientation as proto-oncogene transcription. In this illustration, proto-oncogene transcription is initiated from the 3'LTR of the provirus.

B. Enhancer activation – in this illustration the provirus has inserted upstream in reverse orientation with respect to proto-oncogene transcription. Enhancer elements contained within the U3 region of the provirus activate gene expression from the endogenous promoter. This mechanism of insertional mutagenesis can occur when the provirus has integrated upstream or downstream of a proto-oncogene in either transcriptional orientation.

C. Generation of a readthrough transcript – Provirus insertion needs to occur in the same orientation as proto-oncogene transcription. A transcript is initiated from the 5'LTR of the provirus and reads through into downstream cellular sequences. Splicing generates a transcript in which the viral leader sequence is fused to exons of a cellular proto-oncogene.

D. Disruption of the 3'UTR – Provirus insertion in 3'UTR AU-rich elements (AREs) that regulate the stability of mRNA, can result in transcripts with a prolonged half-life.

(Reproduced from Chapter 10 in Retroviruses, 1997, JM Coffin, SH Hughes, HE Varmus with permission from Cold Spring Harbor Laboratory Press).

1.2.2.6. Identification of CIS in Mo-MLV induced haematopoietic malignancies

Analysis of retroviral integration sites (RIS) in tumours induced by chronically transforming retroviruses revealed many of these had occurred in the same loci. Loci that were found to be frequent targets of retroviral integration in independent tumours are so called common insertion sites (CIS). The genes near some CIS, were known proto-oncogenes, whose sequence was present as v-onc in acutely transforming retroviruses. For example, c-myc and c-myb (discussed below) are transduced by acutely transforming retroviruses of chickens. Analysis of genes near other CIS led to the identification of previously unknown proto-oncogenes, e.g. the serine threonine kinase Pim-1 (Cuypers *et al.* 1984).

The most commonly found RIS in T cell lymphomas following neonatal Mo-MLV infection of C57BL and BALB/c mice, were in c-myc (40%), N-myc (35%) and Pim-1 (45%) (Reviewed in (Jonkers and Berns 1996)). In a quarter of tumours insertions in both c-myc and Pim-1 were found. Most T cell lymphomas in these mice were oligoclonal with RIS in both c-myc and Pim-1 present in independent clones. The acquisition of new RIS during tumour progression was frequently observed after serial transplantation of tumours in syngeneic mice (Cuypers *et al.* 1986). In these tumours, c-myc expression occurred mostly by enhancer activation; whereas N-myc and Pim-1 expression was activated by disruption of a 3'UTR mRNA destabilising element. MLV infection of adult pristane treated BALB/c induces myeloid leukaemia (Shen-Ong and Wolff 1987). In all tumours, retroviral insertion sites in c-myb were found, resulting in expression of either an amino or carboxy terminal truncated c-myb protein.

In recent years, the use of retroviral insertional mutagenesis screens has widened beyond the identification of new oncogenes. Retroviral insertional mutagenesis can accelerate tumourigenesis in tumour prone mouse models (Berns *et al.* 1989). Analysis of CIS in these tumours allows identification of cooperating oncogenes. Tumour progression genes were recently identified by this approach in the CD2 myc/Runx2 model, that already expresses two oncogenes predisposing it to lymphoma (Stewart *et al.* 2007). Use of the $Blm^{-/-}$

model, which has a high rate of loss of heterozygosity, allowed the identification of tumour suppression genes by insertional mutagenesis (Suzuki *et al.* 2006). Inactivation of tumour suppressor genes by insertional mutagenesis is rare, as it requires either the production of a dominant negative protein from the disrupted allele, or inactivation of both alleles. Loss of both p53 alleles is a common occurrence in Fr-MLV induced erythroleukaemias (Johnson and Benchimol 1992). SFFV insertions in one of these p53 alleles were detected in some tumours, and are thought to contribute to p53 loss in these cells (Ben-David *et al.* 1990).

The Retrovirus Tagged Cancer Gene Database (RTCGD) currently holds some 500 CIS from studies using 28 different mouse models (Akagi *et al.* 2004). This database was consulted to find out whether any RIS integration sites in my IL-3 independent mutants of the BAF3/BCL15 cell line were known CIS. The BCL15 cell line we used, already expressed the human Bcl2 oncogene. An interaction search of the RTCGD database was performed to identify genes found in the same tumours as Bcl-2. Nine such genes were identified from 4 different tumour models; they are: growth factor independent 1 (Gfi1), wingless-related MMTV integration site 6 (Wnt6), SET translocation (Set), RAR-related orphan receptor gamma (Rorc), Harvey rat sarcoma virus oncogene 1 (Hras), transcription factor 12 (Tcf12), v-myc myelocytomatosis viral related oncogene (Mycn), Myc and Pim-1. c-myc was shown to collaborate with Bcl-2 in transforming BAF3 cells to IL-3 independence (Malde and Collins 1994).

1.2.3. Lentiviruses with emphasis on HIV

1.2.3.1. Introduction

Acquired Immunodeficiency Syndrome (AIDS) as a clinical entity was first recognised in 1981. Pneumocystis carinii pneumonia (PCP), until then only seen in severely immunocompromised individuals, was diagnosed in 5 previously healthy homosexual males in Los Angeles. The virus that causes AIDS was isolated in 1983 (Barre-Sinoussi *et al.* 1983) and was subsequently named Human Immunodeficiency Virus (HIV). In 1985 it became clear that "slim disease" that had become increasingly prevalent in parts of Sub Saharan Africa was caused by the same virus (Serwadda *et al.* 1985). An estimated 25 million people worldwide have died of HIV/AIDS since 1981. This year marks the 25th anniversary of the discovery of HIV. The introduction of highly active anti-retroviral therapy (HAART) in 1996, led to a dramatic decrease in mortality and morbidity of HIV-infected individuals in the developed world. However, as yet no permanent cure or effective vaccine for HIV/AIDS has been found.

1.2.3.2. Origins of HIV

The distribution of HIV-1 is global. In 1986 a related virus, HIV-2, was isolated, which is prevalent mostly in West Africa. HIV-2 also causes AIDS. However, the onset of disease is slower compared to HIV-1 and the disease is less severe. HIV-1 and HIV-2 are phylogenetically related to simian lentiviruses from chimpanzees (SIZcpz) and sooty mangabeys (SIZsm), respectively. At least 3 cross species transmissions of SIZcpz from chimpanzees to humans are thought to have occurred in Western Central Africa (Cameroon/Gabon), resulting in the establishment of HIV-1 groups M, O and N (Reviewed in (Heeney *et al.* 2006)). SIZcpz and SIVsm are non-pathogenic in their natural hosts; although SIVmac, which is closely related to SIVsm, causes AIDS in macaques.

1.2.3.3. Clinical course of HIV infection

The main route of HIV transmission is sexual. Upon HIV infection, people may experience a mononucleosis-like seroconversion illness. Opportunistic infections may occur at this time as a result of acute immunosuppression. HIV infection is then usually asymptomatic for many years. During this time there is a gradual decline in $CD4^+$ T cell count. Symptomatic disease usually presents itself when a patient's $CD4^+$ T cell count is less than 500 per µl. At this time, patients become increasingly susceptible to opportunistic infections. AIDS is a clinical syndrome defined by 20 AIDS-defining illnesses: opportunistic infections such as PCP and/or AIDS defining malignancies such as Kaposi sarcoma (KS), Non-Hodgkin's lymphoma (NHL) and anogenital cancers.

1.2.3.4. HIV immunopathogenesis

HIV infection causes a specific immunodeficiency of $CD4^+$ helper T cells. A long held view was that this is the result of gradual virus mediated death of $CD4^+$ T cells during the asymptomatic stage of infection. However, recent studies suggest most of the damage to the immune system is done during acute infection (Reviewed in: (Haase 2005) (Picker 2006) (Grossman *et al.* 2006)).

Establishment of HIV infection in most cases occurs in the mucosal immune system in the lamina propria beneath the mucosal epithelial cell surface of the urogenital tract. Relatively few activated $CD4^+$ T cells, in which HIV replicates most efficiently, are present in the mucosal immune system. The most abundant cell type are $CD4^+$ effector memory T cells, which express CCR5 making them readily infectable by transmitted M-tropic HIV-1 (Mattapallil *et al.* 2005). These cells are "resting" but do support HIV replication, and are crucial in the establishment of HIV infection because of their abundance. Dendritic cells (DC) are also found in mucosal epithelia and are thought to be an important target early in HIV-1 infection (Reviewed in: (Piguet and Steinman 2007)). DC are less readily infected than $CD4^+$ T cells. However, they can capture HIV-1 virions via C-type lectins e.g. DC-SIGN on their cell surface, without necessarily becoming infected themselves. These virions are sequestered in intracellular compartments

within DC without being degraded, so they remain infectious. After migration to the lymph nodes, DC very efficiently transfer these virions to $CD4^+$ T cells via direct cell-to-cell contact (McDonald *et al.* 2003; Piguet and Sattentau 2004). This is thought to greatly facilitate transmission at this site.

Once infection is established in the draining lymph node, it rapidly disseminates to other lymphoid tissue compartments such as the peripheral lymph nodes, the spleen and gut-associated lymphoid tissue (GALT). 60% of total $CD4^+$ T cells, mostly effector memory T cells, are found in the GALT and this cell population is almost completely destroyed early in infection by direct cytopathic effects of the virus (Li *et al.* 2005; Mattapallil *et al.* 2005).

The GALT is also instrumental in raising an immune response against HIV. Neutralising antibodies and cytotoxic T cell response against HIV have been detected, but due to its high mutation rate HIV quickly evolves to evade both.

It is now believed that the reason why there is a drop in viraemia after acute infection is because the supply of $CD4^+$ T cells in the gut is exhausted; not because a potent immune response is raised against the virus (Phillips 1996). This pool is somewhat replenished by naïve and central memory T cells that have not been targeted in acute infection. This process of replenishment is stimulated by a generalised state of immune activation which exists at the time of clinical latency (Grossman *et al.* 2006). Immune activation also increases the number of activated T cells, providing a source for ongoing HIV replication. It is thought that the loss of this ability to regenerate effector memory T cells is responsible for the eventual drop in overall $CD4^+$ T cell numbers. This is seen clinically by an increased susceptibility to opportunistic pathogens.

Activated $CD4^+$ T cells survive for only a few days following infection. The introduction of HAART allowed the study of HIV reservoirs in patients (Simon and Ho 2003). It became evident that latent reservoirs are established very early during infection. Long lived cell populations such as truly resting/quiescent memory $CD4^+$ T cells and dendritic cells contain transcriptionally inactive proviruses. Upon reactivation from latency, virus is produced by these cells. This explains the need for life long antiretroviral therapy to suppress HIV replication.

1.2.3.5. HIV associated malignancies

HIV-infected patients are more prone to a variety of cancers. The factors involved in this are thought to be chronic immune stimulation, impaired tumour surveillance and infection with oncogenic herpes- or papilloma viruses. The latter are associated with three AIDS defining malignancies: Kaposi sarcoma, Non Hodgkin's lymphoma and anogenital cancer (Boshoff and Weiss 2002).

1.2.3.5.1. Kaposi sarcoma (KS)

This is the most common cancer in AIDS patients. KS is a malignancy of the lymphatic endothelial precursor cells caused by Human Herpes Virus 8 (HHV-8) also known as Kaposi sarcoma herpes virus (KSHV). Organs in which KS lesions are commonly found are the skin, mouth, gastrointestinal tract and lungs.

1.2.3.5.2. Non Hodgkin's Lymphoma (NHL) (Killebrew and Shiramizu 2004)

AIDS-associated NHLs are almost uniquely B cell in origin. 80% of these are systemic lymphomas. Of these, only a minority (20%) are confined to the lymph nodes. Half of them are associated with Epstein Barr Virus (EBV) infection. The other half is thought to result from oligoclonal B cell expansion in response to chronic antigen stimulation. In the absence of $CD4^+$ T cell control over B cell proliferation, a population emerges that is at increased risk of transformation. Primary CNS lymphomas make up about 20% of NHL. They are always associated with EBV infection. KSHV also infects B cells and is able to transform these cells in a manner similar to EBV. KSHV is associated with rare NHL, namely primary effusion lymphoma (PEL) and plasmablastic lymphoma, a variant of Multicentric Castleman's disease.

1.2.3.5.3. Anogenital and Cervical Cancers

HIV infected people are at increased risk of cancers caused by human papilloma viruses (HPV). These include cervical cancer in women and anorectal and genital cancer in both sexes.

1.2.3.5.4. HIV-induced insertional mutagenesis in rare lymphomas

Most of the target cells for HIV infection in vivo are terminally differentiated cells. In gene therapy, pseudotyping with heterologous envelope proteins confers a broader tropism to lentiviral vectors. The targets of lentiviral vector transduction are likely to be less mature progenitor cells such as $CD34^+$ haematopoietic stem cells (HSC). These cells may be more susceptible to transformation. Although $CD34^+$ cells can be infected by wild type HIV-1 ex vivo, this has rarely been detected in HIV infected patients (Moses *et al.* 1998). Has direct involvement of HIV in transformation of its targets cells ever been demonstrated in HIV infected patients? A case report exists that describes a rare $CD4^+$ T cell lymphoma found in an AIDS patient (Herndier *et al.* 1992). This lymphoma contained a monoclonal HIV-1 provirus integration 1 kb upstream of the c-fes/fps proto-oncogene (Shiramizu *et al.* 1994). c-fes/fps expression was activated in tumour cells. Intriguingly, the same group found monoclonal HIV-1 integrated in the c-fes/fps locus in HIV infected macrophages in 3 more polyclonal lymphomas of mixed immunophenotype (Shiramizu *et al.* 1994).

1.3. Vector construction

In the context of gene delivery, vectors are vehicles used to introduce genes into target cells. Both non-viral and viral vectors exist. Table 1.3. summarises the properties of the different viral vectors that are currently being developed for use in gene delivery. This section shall focus on the development of vectors from simple (MLV) and complex (HIV-1) retroviruses. Initially, features particular to these two types of vectors will be discussed separately. Common concepts such as insert configuration, self-inactivating vector design, production and pseudotyping will then be addressed.

1.3.1. Retroviral vectors

Retroviruses integrate into the genome of host cells they infect. This property allows stable gene expression, which is highly desirable in the context of gene therapy. It also allows a transgene introduced into a progenitor cells to be passed on to all its progeny. However, the introduction of foreign DNA into the cell genome per se is potentially mutagenic. The problem is confounded by the presence of potent gene regulatory elements in the LTRs. These elements are retained in retrovirus derived vectors.

Acutely transforming retroviruses are naturally occurring vectors that transduce oncogenes into cells. This observation led to the eventual exploitation of retroviruses as vectors for gene therapy. Acutely transforming retroviruses were mostly replication defective, due to extensive deletions in their viral coding regions, and relied on viral proteins provided by replication competent helper viruses to replicate. Two reviews on retroviral vector design were consulted: (Takeuchi and Pizzato 2000; Buchschacher 2001)

Table 1.3. The main groups of viral vectors

Vector	Genetic Material	Packaging Capacity	Tropism	Inflam-matory potential	Vector genome forms	Main limitations	Main advantages
Enveloped							
Retrovirus	RNA	8 kb	Dividing cells only	Low	Integrated	Only transduces dividing cells; integration might induce oncogenesis in some applications	Persistent gene transfer in dividing cells
Lentivirus	RNA	8 kb	Broad	Low	Integrated	Integration might induce oncogenesis in some applications	Persistent gene transfer in most tissues
Herpes Simplex Virus -1	dsDNA	40 kb * 150 kb **	Strong for neurons	High	Episomal	Inflammatory; transient transgene expression in cells other than neurons	Large packaging capacity; strong tropism for neurons
Non-enveloped							
AAV	ssDNA	< 5 kb	Broad, with the possible exception of haematopoietic cells	Low	Episomal (> 90%) Integrated (< 10%)	Small packaging capacity	Non-inflammatory; non pathogenic
Adenovirus	dsDNA	8 kb * 30 kb §	Broad	High	Episomal	Capsid mediates potent inflammatory response	Extremely efficient transduction of most tissues

* Replication defective ** Amplicon § Helper dependent

Adapted with permission from Macmillan Publishers Ltd: Thomas *et al* Nature Review Genetics **4**: 346 – 358, 2003.

1.3.1.1. Retroviral transfer vectors

The basic principle behind the design of a replication defective retroviral transfer vector is to replace the viral coding regions (gag, pol, env) by the transgene that is to be delivered to target cells. This is facilitated by the fact that the viral coding regions are found in the centre of the genome separated from the so called *cis*-acting elements. The latter are retained in the transfer vector. They include the LTRs, necessary for transcription initiation, polyadenylation and reverse transcription; the packaging (Ψ) signal, required for transfer vector incorporation into particles; and the PBS and PPT, necessary for, respectively, first and second strand synthesis during reverse transcription. Most retroviral transfer vectors currently in use were derived from the genomes of ecotropic MLVs, particularly Mo-MLV. These vectors can hold inserts up to 7kb in size.

1.3.1.2. Retroviral packaging constructs

To enable vector virus production, proteins encoded by the viral gag, pol and env genes must be provided in *trans*. They are provided by retroviral packaging constructs. Packaging constructs have undergone a number of modifications over the years. The simplest such constructs were retroviral genomes from which the Ψ signal was deleted, such as in the 1st generation Ψ-2 packaging cell line (Mann *et al.* 1983). However, because the high degree of homology between transfer and packaging constructs, only one recombination event was required to generate replication competent retrovirus (RCR), which is of course highly undesirable. In 2nd generation constructs, in addition to the Ψ signal, the PPT and 3'LTR were removed (Miller and Buttimore 1986). This left just the 5'LTR driving gag, pol and env expression. In 3rd generation "split" packaging constructs, such as those in Ψ CRE and Ψ CRIP (Danos and Mulligan 1988), the gag, pol genes were separated from the env gene. Three recombination events would now be needed to generate RCR. Env expression from a separate construct also allowed pseudotyping with heterologous env proteins (See Section 1.3.6.)

1.3.2. Lentiviral vectors

Vectors derived from gammaretroviruses can only transduce dividing cells. In recent clinical gene therapy trials using these vectors, haematopoietic stem cells (HSC), which are mostly non-dividing, required stimulation with cytokines prior to transduction. This necessitated a 4 day ex vivo culture time of HSC (See Table 1.4. in the next section). This can be shortened when lentiviral vectors are used, because they efficiently transduce non-dividing as well as dividing cells. Non-dividing target cells such as neurons, hepatocytes and myocytes can be transduced, widening the application of lentiviral vectors. Lentiviral vectors based on the genomes of a simian (SIV), feline (FIV) and equine (EIAV) lentiviruses have now been designed, but those based on HIV-1 are currently at the most advanced state of development. The following reviews on lentiviral vectors were consulted: (Woods *et al.* 2002; Delenda 2004)

1.3.2.1. Lentiviral transfer vectors

The transfer vector contains the only HIV-1 sequence transferred to the target cell. The "first generation" lentiviral transfer vector pHR' contains all HIV-1 cis-acting sequences necessary for packaging (Ψ), reverse transcription (PBS, PPT, LTRs) and integration (integrase attachment sites)(Naldini *et al.* 1996). The RRE, part of the env ORF, is also retained. Additional cis-acting sequences were later introduced into lentiviral transfer vectors. The cPPT, located in the central region of lentiviral genomes, enhances nuclear translocation of pre-integration complexes and increases transduction efficiency in both dividing and non-dividing cells (Zennou *et al.* 2000). The post-transcriptional regulatory sequence of Woodchuck hepatitis B virus (WPRE) improves transgene expression levels when inserted downstream of the transgene (Zufferey *et al.* 1999).

1.3.2.2. Lentiviral packaging constructs

The lentiviral packaging systems described are optimised for transient transfection (See Section 1.3.5). The "first generation" lentiviral packaging construct (pCMVΔR9) encodes all HIV-1 genes with the exception of *env* and the accessory gene *Vpu* (Naldini *et al.* 1996). The HIV-1 packaging signal is deleted

in this construct; the HIV-1 5' and 3'LTR are replaced by a heterologous CMV promoter and a polyadenylation signal from the human insulin gene, respectively. "Second generation" lentiviral packaging constructs e.g. pCMVΔ8.91 only express gag, pol and the regulatory genes tat and rev, but none of the HIV-1 accessory genes (Zufferey *et al.* 1997). Tat is required during vector production for transcription from the HIV-1 5'LTR of the transfer vector genome; Rev is necessary during production for nuclear export of transfer vector genomic RNA and gag-pol mRNA. Gag-pol mRNAs are targeted for degradation in the absence of Rev, because their coding regions contain so called inhibitory sequences (INS). "Third generation" lentiviral packaging contructs such as pMDL g/p RRE encode only gag, pol and the RRE (Dull *et al.* 1998). Rev is expressed from a separate construct. Tat is dispensable in combination with lentiviral transfer vectors, in which the 5'LTR U3 region has been replaced by a heterologous promoter that drives transfer vector RNA production in producer cells. However, Rev continues to be an absolute requirement for transfer vector RNA nuclear export. Gag-pol codon optimisation inactivates gag-pol INS, making it Rev-independent.

1.3.3.Insert configuration for transgene expression

In the simplest retroviral vectors, the transgene is expressed from the 5'LTR. In first generation "intronless" retroviral vectors, the transgene start codon is in place of the Gag ATG (Eglitis *et al.* 1985). In the MFG vector, the retroviral SD and SA are present; the transgene start codon is in the same position as that of the deleted env gene (Riviere *et al.* 1995). Increased levels of spliced RNA in transduced cells explains higher transgene expression by these vectors (Krall *et al.* 1996). As Tat is not present in target cells, in lentiviral vectors a heterologous promoter, which is active in the target cell, is always required for transgene expression. This is also the case with self-inactivating vectors (See section 1.3.4.). The choice of heterologous promoters can control transgene expression: Constitutively active promoters derived from retroviral LTRs (such as spleen focus forming virus (SFFV), myeloproliferative sarcoma virus (MPSV)); other promoters from other viruses such as cytomegalovirus (CMV) or cellular genes

(Phospho glycerate kinase (PGK), elongation factor 1α (EF1α)) can express transgenes in a range of cell types. The use of tissue-specific promoters potentially improves vector safety and efficacy by limiting transgene expression to specific target cells only. For example, to avoid an immune response to Factor IX, a liver-specific albumin promoter was used in one study (Follenzi *et al.* 2004). Similarly, the incorporation of microRNA (miRNA) target sequences in the transgene transcript prevents transgene expression only in cell types in which the particular miRNA is expressed (Brown *et al.* 2007). In this way transgene expression in antigen presenting cells can be avoided. Regulation of transgene expression is sometimes desirable; this can be achieved by using inducible promoters e.g. those controlled by tetracycline (Paulus *et al.* 1996). More than one transgene can be expressed from a single vector. This can be done using biscistronic vectors, where a second transgene is expressed downstream of an internal ribosomal entry site (IRES) (Adam *et al.* 1991). More efficient expression of the second transgene can be achieved using dual promoter constructs.

1.3.4. Generation of self-inactivating (SIN) vectors

In SIN vectors, the enhancer and promoter elements contained within the U3 region are removed from the 3'LTR of the transfer vector construct. Following reverse transcription, these gene regulatory elements will have been deleted from both LTRs. SIN vectors are believed to be safer for the following reasons:
the number of enhancer/promoter elements in these vectors is reduced to just the internal heterologous promoter driving transgene expression. The risk of insertional mutagenesis by SIN vectors is therefore theoretically lower. Secondly, transgene expression may be more efficient, due to lack of transcriptional interference between the LTRs and the internal promoter. Minimal packaging systems have been developed for the production of both retro- and lentiviral vectors; SIN vector design yet further reduces the chance of RCR generation.
Lastly, SIN lentiviral vectors are safer in the clinical setting. In case patients treated with these vectors become infected with wild-type HIV-1, the risk of SIN vector mobilisation by HIV-1 is lower.

Enhancer/promoter deleted retroviral SIN vectors were first described in the late 1980s (Yu *et al.* 1986; Yee *et al.* 1987). In MLV, upstream sequence elements (USE) that enhances polyadenylation are dispersed throughout the U3 region. Extensive U3 deletion in SIN vectors leads to inefficient polyadenylation at the 3'LTR, increasing the risk of read-through into downstream cellular genes. This can be overcome by introducing heterologous USE in the deleted U3 region in the 3'LTR of SIN vectors (Schambach *et al.* 2007) or by placing a WPRE downstream of the transgene (Higashimoto *et al.* 2007).

In the first lentiviral SIN vector pHR SIN, 400bp of the U3 region were removed from the HIV-1 3'LTR of the lentiviral transfer vector plasmid (Zufferey *et al.* 1998). In HIV-1, the USE that enhances polyadenylation is found downstream of the TATA box, just upstream of the R region (Valsamakis *et al.* 1991).This allowed more extensive deletion of U3 sequence for the purpose of SIN lentiviral vector generation. In pHR SIN, only the first 35bp of U3, containing the integrase attachment site, and the final 20bp of U3 containing the USE essential for polyadenylation at the 3'LTR, are left. The core promoter, basal enhancer and the bulk of the modulatory region have all been removed (See Figure 1.5.).

When analysed by Northern blot, no transcripts originating from the SIN LTR could be detected (Zufferey *et al.* 1998), whereas some transcripts from a wild type LTR could be detected in HeLa cells, even in the absence of Tat. However, SIN lentiviral vectors are not completely transcriptionally silent (Logan *et al.* 2004). It was found that the DBF-1 and Sp1 transcription factor binding sites present in the leader region downstream of the transcription start site, could induce transcription from the U3/R junction, resulting in some residual promoter activity, about 15% that of a wild type LTR. Mutation of the DBF-1 and Sp1 sites could reduce this even further, though mutation of the latter site resulted in a dramatic drop in titre, because the downstream Sp1 site overlaps with the lentiviral packaging signal and led to reduced full length genomic transcript packaging (Logan *et al.* 2004).

1.3.5. Production

High titre retro- and lentiviral vector particles can be produced by transient transfection (Soneoka *et al.* 1995; Naldini *et al.* 1996). This involves transfecting transfer vector, packaging and envelope coding plasmid constructs into highly transfectable cells, most commonly 293T cells. SIN lentiviral vectors can be produced by transient transfection to similar titres as non SIN vectors, whereas the titres of SIN retroviral vectors are usually several fold lower than that of their wild type counterparts.

For clinical applications, vector production using stable packaging cells is desirable. These cells can be extensively screened for RCR generation, allowing the production of large batches of quality controlled vector. In recent clinical trials, the retroviral packaging cell lines Ψ CRIP (Danos and Mulligan 1988), Gp+ AM12 (Markowitz *et al.* 1988) and PG13 (Miller *et al.* 1991) were used (See Table 1.5.). The latter produces Gibbon Ape Leukaemia Virus (GALV) pseudotyped particles; the former two produce MLV-A pseudotyped particles. All three cell lines are derived from mouse cells; retroviral vectors produced from these cells were used for ex vivo transduction of HSC. They would not be suitable for human in vivo use, as they are rapidly inactivated by human serum (Takeuchi *et al.* 1996). For this reason, retroviral packaging cells derived from human cells were developed (Cosset *et al.* 1995). There is also a reduced risk of cross-packaging endogenous retroviral sequences when human cells are used (Patience *et al.* 1998).

Stable lentiviral packaging cell lines have also been created. In 293T/HT1080-based STAR cells, codon-optimised HIV-1 gag-pol (lacking RRE), Tat and Rev are each expressed from a separate MLV-based retroviral vector (Ikeda *et al.* 2003). STAR cells produce lentiviral particles pseudotyped with either MLV-A, RD114 or GALV. Non-SIN lentiviral transfer vectors can be introduced into these cells by transduction; whereas SIN lentiviral transfer vectors can only be introduced by transfection.

1.3.6. Pseudotyping

The following reviews on pseudotyping were consulted: (Sanders 2002; Verhoeyen and Cosset 2004). As was alluded to in Section 1.3.1., the introduction of packaging constructs in which the gag-pol coding regions are separated from those of env, allowed the incorporation of heterologous envelope glycoproteins into retroviral particles. Vector particles with an altered tropism from that of the retroviral vector genome could be generated in this way. Retroviral vectors based on the genome of ecotropic MLV (Mo-MLV) are commonly pseudotyped with MLV-A envelope, to broaden tropism to include human cells. The gammaretroviral envelope proteins from RD114 and GALV can be used to pseudotype retroviral but not lentiviral cores. Cytoplasmic tail modifications of the RD114 and GALV envelope glycoproteins, for example by replacing it with the tail of MLV-A which efficiently pseudotypes lentiviral cores, overcomes this limitation (Verhoeyen and Cosset 2004). Retroviral and lentiviral particles can also be pseudotyped with envelope proteins from other viruses such as influenza haemagglutinin (Hatziioannou *et al.* 1998), Ebola (Wool-Lewis and Bates 1998), hepatitis C virus (Bartosch *et al.* 2003). The vesicular stomatitis virus G protein (VSV-G) is commonly used (Burns *et al.* 1993; Naldini *et al.* 1996). VSV-G confers a broad tropism onto pseudotyped particles; it also allows concentration by ultracentrifugation enabling high titre vector preparation. However, VSV-G pseudotyped vector cannot be used in vivo, as it is rapidly inactivated by human serum. Pseudotransduction (the transfer of transgene protein product in the absence of stable integration) has also been documented. Production of VSV-G pseudotyped vector particles from stable cell lines is problematic as stable expression of VSV-G protein is toxic to cells. Inducible VSV-G expression has therefore been required (Sanders 2002). Generally, the titres produced by these stable lines are lower than those achieved by transient transfection.

Envelope proteins have also been modified with the purpose of targeting vector particles to specific cell types. Viral envelope glycoproteins were engineered to contain ligands, peptides or single chain antibodies (scFc) (Hatziioannou *et al.* 1999). Though efficient binding to specific target cells could be demonstrated, this was not accompanied by efficient virus entry. It seems the fusogenic ability of the modified envelope glycoproteins had been compromised, resulting in low titre vector (Verhoeyen and Cosset 2004). In a recent paper, separate binding and fusogenic molecules were incorporated into the envelope of lentiviral particles (Yang *et al.* 2006). The binding molecule recognises a specific cell surface receptor; binding induces endocytosis of vector particle. The fusogenic molecules were chosen to induce fusion with the endosomal membrane at low pH. Lentiviral particles expressing an engineered glycoprotein from Sindbis virus could target dendritic cells in vivo by recognising DC-SIGN (Yang *et al.* 2008).

1.4. Gene Therapy

1.4.1. Introduction

Gene therapy can be defined as the introduction of genetic material (DNA or RNA) into cells for therapeutic effect. In recent years, gene therapy has been explored as a treatment for diseases with limited alternative therapeutic options.

The potential applications for gene therapy are numerous. Worldwide, gene therapy clinical trials are most commonly performed for cancer (66.5%). This is followed by cardiovascular diseases (9.1%), monogenic inherited disorders (8.3%) and infectious diseases (6.5%) (Edelstein *et al.* 2007). In just under a quarter of clinical trials, retroviral vectors were used for gene delivery. This section shall focus on the use of retroviral vectors in gene therapy. The therapeutic benefit of gene therapy using these vectors was most demonstrable in the treatment of individuals suffering from primary immunodeficiencies. It was in two of these trials that insertional mutagenesis caused malignant cell transformation, which inspired investigation in this thesis.

1.4.2. Retroviral and lentiviral vectors in human gene therapy

One of the earliest uses of retroviral vectors in man was in gene marking studies. To evaluate the efficacy of tumour infiltrating lymphocyte (TIL) treatment in metastatic melanoma, cells were transduced with a retroviral vector so that they could be tracked (Rosenberg *et al.* 1990). Gene transfer to haematopoietic stem cells (HSC) was first performed in cancer patients who needed autologous HSC transplantation (HSCT) as part of their treatment; HSC were transduced with retroviral vectors encoding multidrug resistance (MDR) genes, allowing higher doses of chemotherapy to be administered (Hanania *et al.* 1996). Lymphocytes and HSC are convenient targets for retroviral vector transduction as they are accessible and techniques for ex vivo culture exist. This allows ex vivo transduction of these cells; followed by re-infusion into patients. HSC gene therapy shall be discussed in the next paragraph. Progress in T cell gene therapy has also been made. In an extension of the TIL approach just described, gene transfer was used to create these cells in the first place as opposed to simply marking them (Morgan *et al.* 2006). Autologous PBL were transduced with a retroviral vector encoding a T cell receptor that recognises a melanoma tumour-associated antigen. 15 patients were treated in this study; in 2, transduced TIL could still be detected at 1 year post-infusion. This was accompanied by marked regression of metastatic melanoma lesions. Recently, the first clinical gene therapy trial using a lentiviral vector was reported (Levine *et al.* 2006). Autologous $CD4^+$ T cells of 5 subjects suffering from drug resistant HIV-1, were transduced with a lentiviral vector encoding an antisense gene against the HIV-1 envelope. Transcription of the env antisense gene is from a non-SIN HIV-1 LTR in the vector. Its expression depends on infection of the cell with wild type HIV-1, which provides Tat. Production of progeny of wild-type HIV-1 virions from this cell should be abolished, if degradation of env transcripts by the antisense RNA is effective. Other RNA-based gene therapy approaches, involving transduction of PBL or HSC, to create an HIV-1 resistant pool of lymphocytes are currently being explored (Reviewed in (Rossi *et al.* 2007)).

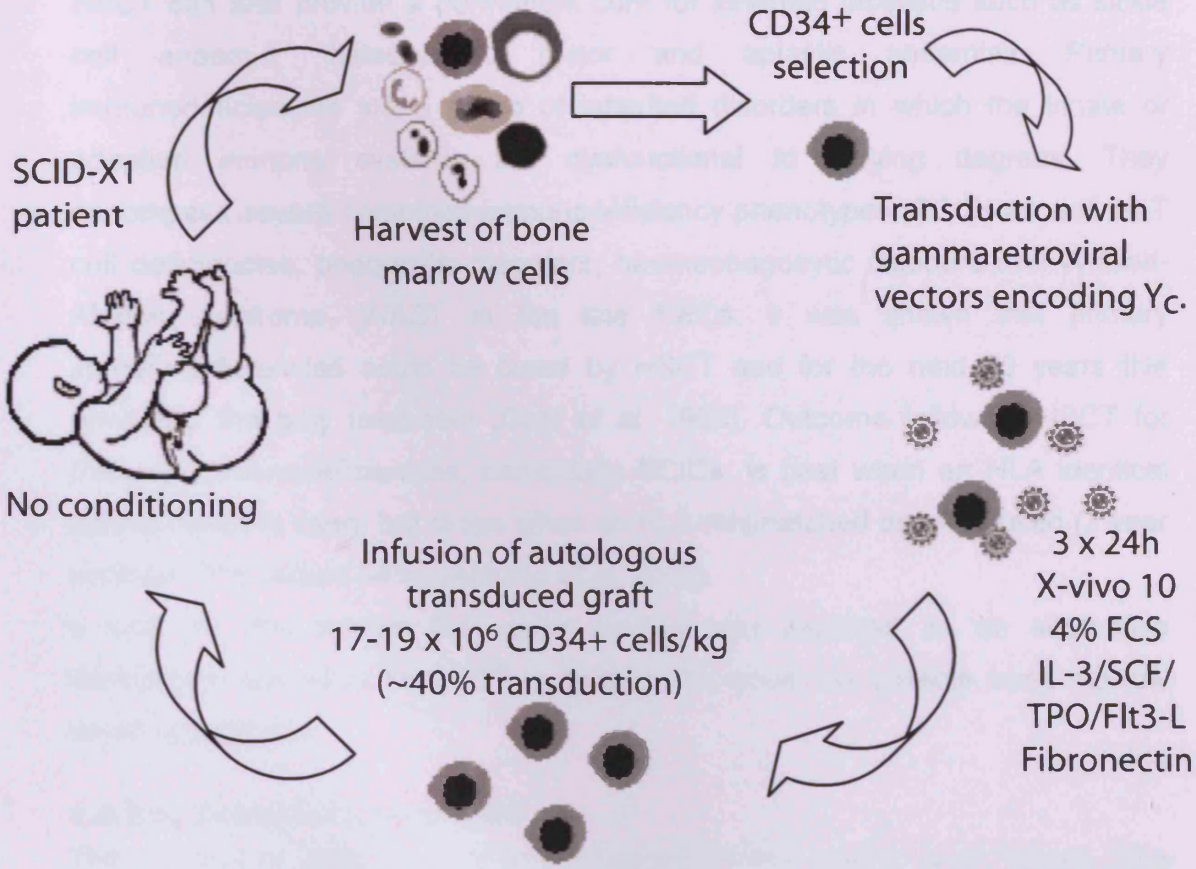


Figure 1.7. Principle of gene therapy to the haematopoietic system
This figure illustrates the gene therapy protocol for the treatment of SCID-X1 used in the trial by Cavazzana-Calvo *et al.* Bone marrow was harvested from affected boys; $CD34^+$ cells were isolated from the bone marrow; these cells were transduced ex vivo with an MFG vector carrying a correct copy of the gamma c transgene; transduced cells were re-infused back into the patient. This resulted in successful reconstitution of the immune system in 9 out of 10 patients treated in this trial.

(Reproduced from Bueren *et al*, Archives of Medical Research **34**(6): 589-99, 2003, with permission from the publisher Elsevier.)

1.4.3. Haematopoietic system gene therapy

Haematological malignancies are the commonest indication for allogeneic HSCT. HSCT can also provide a permanent cure for inherited diseases such as sickle cell anaemia, thalassaemia major and aplastic anaemias. Primary immunodeficiencies are a group of inherited disorders in which the innate or adaptive immune systems are dysfunctional to varying degrees. They encompass severe combined immunodeficiency phenotypes (SCID) as well as T cell deficiencies, phagocytic disorders, haemophagocytic disoders and Wiskott-Aldrich syndrome (WAS). In the late 1960s, it was shown that primary immunodeficiencies could be cured by HSCT and for the next 30 years this remained the only treatment (Gatti *et al.* 1968). Outcome following HSCT for primary immunodeficiencies, particularly SCIDs, is best when an HLA identical sibling donor is used, but drops when an HLA mismatched donor is used (3 year survival 77% versus 54%) (Antoine *et al.* 2003).

It was for this reason that gene therapy was explored as an alternative therapeutic approach to HSCT in patients for whom no suitable bone marrow donor is available.

1.4.3.1. Transduction of HSC

The concept of gene therapy to the haematopoietic system is as follows (See Figure 1.7.): To begin with a patient's own HSC are used – these can be derived from bone marrow, cord blood or peripheral blood following G-CSF mobilisation. $CD34^+$ cells are fractionated from total bone marrow or peripheral blood cells, these are a heterogenous population of haematopoietic progenitor cells and only as little as 0.2% are HSC (Bhatia *et al.* 1997). The hallmark properties of HSC are self-renewal and the potential to develop into mature blood cells of different lineages. In clinical gene therapy protocols so far, the therapeutic gene was introduced by means of a retroviral vector based on MLV. Since HSC are quiescent and retroviral vectors can only transduce dividing cells, it is necessary to pre-stimulate $CD34^+$ cells ex vivo for 1-2 days in the presence of cytokines prior to vector transduction. (As pointed out previously, this ex vivo culture time

can be shortened when lentiviral vectors are used.) To enhance vector transduction of HSC, they are grown in the presence of vector supernatant in vessels coated with recombinant fibronectin (Pollok and Williams 1999). Gene therapy corrected autologous HSC are then infused back into the patient, who depending on the underlying condition has or has not received pre-conditioning chemotherapy to allow better engraftment.

1.4.3.2. Clinical gene therapy trials to treat primary immuno-deficiencies

Details of recent clinical gene therapy trials for primary immunodeficiencies are shown in Table 1.4.

In March 1999, what turned out to be the first successful human gene therapy trial was initiated in France for the treatment of X-linked SCID (X-SCID) (Cavazzana-Calvo *et al.* 2000). In this condition, the common γ chain gene is faulty; the protein encoded by this gene is a component of the IL – 2/4/7/9/15/21 receptors. In X-SCID patients, a failure of IL-7 and IL-15 signalling in HSC results in the complete absence of T cells and NK cells, respectively. Though, X-SCID patients have B cells, these are dysfunctional. Unless they are treated by HSCT, X-SCID affected boys succumb to infection within the first year of life. In the French SCID trial, 10 patients were treated. Autologous HSC were transduced ex vivo with an MLV-A pseudotyped MFG retroviral vector expressing γc. In 9 out of 10 children, successful immune reconstitution, particularly of the T and NK cell, but less so of the B cell compartments, was achieved. Similarly favourable outcomes were obtained in a British SCID gene therapy trial (Gaspar *et al.* 2004), in which 10 children were treated. The protocols used in X-SCID trials were very similar, except for the following differences: In the UK trial, the vector was pseudotyped with GALV, $CD34^+$ cells were transduced in serum-free medium containing slightly lower concentrations of IL-3 and lastly the vector used in this trial did not contain the B2 mutation in the PBS. (It was shown recently that TRIM28 binds the Mo-MLV PBS in embryonic carcinoma and stem cells, resulting in transcriptional silencing of the provirus (Wolf and Goff 2007). The B2 mutation prevents TRIM28 binding.) The main difference observed between the

French and UK X-SCID trials is that the rate of T cell reconstitution in the latter trial is slower. Patients did not receive pre-conditioning chemotherapy in either trial. The success of both X-SCID gene therapy trials was largely attributed to the huge selective advantage of gene corrected cells in vivo that followed the gene therapy mediated restoration of their ability to respond to proliferation signals. Subsequent developments in the X-SCID gene therapy trials will be discussed in section 1.4.4.1.

The adenosine deaminase (ADA) enzyme which is deficient in ADA-SCID, is involved in purine metabolism; its deficiency results in abnormal T, B and NK cell development. ADA-SCID can be cured by HSCT; in addition lifelong replacement therapy with polyethylene glycol ADA (PEG-ADA) is available to those for whom no suitable donor can be found. Though PEG-ADA prolongs survival, not all ADA-SCID patients respond equally well to it. ADA-SCID was the first inherited monogenic disorder treated by gene therapy (Blaese *et al.* 1995; Bordignon *et al.* 1995; Kohn *et al.* 1995). However, immune reconstitution in these early gene therapy trials was poor. It is thought that due to the continued use of PEG-ADA in gene therapy treated patients, the potential selective advantage of gene modified cells was abolished. In 2002, the first successful gene therapy trial for ADA-SCID was reported (Aiuti *et al.* 2002). Initially, two children for whom PEG-ADA was unavailable, were treated. They received non-myeloablative conditioning with busulfan prior to re-infusion of gene corrected cells, to allow better engraftment and proliferation of these cells. Full immune reconstitution as well as metabolic correction was achieved in all 5 children treated according to this protocol (Aiuti *et al.* 2007). A similarly encouraging outcome was achieved in the UK using a comparable protocol of elective PEG-ADA discontinuation with mild pre-conditioning (Gaspar *et al.* 2006).

Table 1.4. Overview of recent clinical gene therapy trials for 1° immunodeficiencies

Trial	SCID-X1 France [1, 2, 3]	SCID-X1 UK [4]	SCID-ADA Italy [5]	SCID-ADA UK [6]	CGD [7]
# of patients	10	10	5	1	3
Transgene	γc	γc	ADA	ADA	gp91phox
Vector	MFG (B2) (Mo-MLV)	MFG (Mo-MLV)	GIADA-1 LXSN	SFada WPRE	SF71gp91
Pseudotype	MLV-A	GALV	MLV-A	GALV	GALV
Packaging cell line	ψcrip	PG13	Gp+AM12	PG13	PG13
Target Cell (ex vivo culture time)	BM $CD34^+$ (24+72)	BM $CD34^+$ (40+56)	BM $CD34^+$ (24+72)	BM $CD34^+$ (40+56)	PB $CD34^+$ (36+72)
Conditioning	-	-	Busulfan	L-Busulfan	Melphalan
Leukaemias	4	1	-	-	-

[1] (Cavazzana-Calvo *et al.* 2000), [2] (Hacein-Bey-Abina *et al.* 2002), [3] (Hacein-Bey-Abina *et al.* 2003)
[4] (Gaspar *et al.* 2004), [5] (Aiuti *et al.* 2002), [6] (Gaspar *et al.* 2006), [7] (Ott *et al.* 2006)

The successes achieved using gene therapy for the cure of X-linked and ADA-SCID, very convincingly demonstrated the effectiveness of this treatment for disorders of the lymphoid compartment. In both conditions gene modified HSC are thought to have a selective advantage in engraftment over unmodified cells. This would not be so in the case of chronic granulomatous disease (CGD). CGD is a disorder of the myeloid compartment; CGD macrophages lack anti-microbial activity, predisposing CGD patients to severe bacterial and fungal infections. In the X-linked form of CGD, the protein product of the gp91 phox gene is not expected to confer a selective advantage on gene-modified HSC in a gene therapy setting. Two CGD patients, who were in their thirties, received non-myeloablative conditioning prior to reinfusion of retroviral vector corrected PB-mobilised autologous HSC (Ott *et al.* 2006). Vector transduced cells were found predominantly in the myeloid compartment of both patients resulting in functionally corrected phagocytes and clinical improvement. Though myelopoiesis was initially polyclonal, by 5 months after gene therapy, this had become increasingly oligoclonal. Several progenitor cell clones were found to disproportionately contribute to myelopoiesis, a phenomenon known as clonal dominance. These clones were enriched in integration sites near MDS-Evi1, Prdm16 or Setbp1, which are known proto-oncogenes. Emergence of these dominant clones probably contributed to the therapeutic effect of gene therapy in these two patients. Unlike in the X-SCID trial, neither of these patients developed frank leukaemia. Unfortunately, one patient succumbed to fatal infection 2 years after gene therapy. In this patient's phagocytes, gp91 phox expression was lost due to methylation of the SFFV promoter in the retroviral vector.

1.4.4. Gene Therapy and cancer

1.4.4.1. Leukaemia in clinical gene therapy trials

As was alluded to earlier, severe adverse events in the form of leukaemia occurred in 5 children treated by gene therapy for X-SCID. The first 2 and most well documented cases occurred in the French X-SCID trial, some 2 ½ years after gene modified cells were administered (Hacein-Bey-Abina *et al.* 2003). In

both children, malignant clonal expansion of mature T lymphocytes was observed (1 γδ T cell clone in Pt 4; 3 αβ T cell clones in Pt 5). Malignant T cell clones each contained a single copy of the vector; in both cases the vector integration site was found to be near the promoter of the LIM-domain only 2 (LMO2) gene (In the first intron in Pt 4; 3kb upstream in Pt 5). This gene encodes a transcription factor that is considered to be a central regulator of haematopoiesis, and is a known proto-oncogene. LMO2 expression was activated by enhancer sequences within the LTRs of the MFG vector. LMO2 activation is thought to have been the initiating event in leukaemogenesis, however other chromosomal abnormalities were also found in malignant T cell clones. Details of the other two leukaemia cases in the French X-SCID trial were recently presented at the ESGCT annual meeting in November 2007. These occurred in patients 7 and 10, at 68 and 33 months after gene therapy, respectively. Pt 7 had a malignant clonal expansion of immature T cells. These contained one copy of the vector, which had integrated upstream of the cyclin D2 gene, a known proto-oncogene. A malignant cortical thymocyte clone was found in Pt 10. This clone contained two copies of the vector; one integration site was in intron 1-2 of LMO2; the other was in intron 4-5 of Bmi1. In December 2007, it was announced one patient in the British X-SCID gene therapy trial had also developed leukaemia. LMO2 was overexpressed in this leukaemia; in addition mutations in NOTCH were found (A.J. Thrasher, BSGT annual meeting April 2008).

In all, 5 patients out of a total of 20 patients treated by gene therapy for X-SCID experienced adverse effects that can be ascribed to vector mediated insertional mutagenesis. No such adverse effects have so far been seen in the ADA-SCID trial. It is therefore believed there are specific risk factors associated with X-SCID gene therapy. Restoration of signalling by the interleukin receptors of which γc is a part, especially IL-7 signalling, creates a strong proliferative signal for gene corrected cells, that could contribute to transformation of cells that already contained one oncogenic hit due to vector insertion. This was supported by the finding that in a leukaemia induced by wild type MLV, RIS in both LMO2 and γc were found (Dave *et al.* 2004). This provided evidence for cooperation between

LMO2 and γc genes in tumourigenesis. The role of the γc transgene has been the subject of even more intense debate, following a short report in Nature claiming that the γc transgene itself was potently oncogenic (Woods *et al.* 2006). In this study, $\gamma c^{-/-}$ and normal HSC were transduced at high MOI with a lentiviral vector encoding γc. A third of mice developed T cell lymphoma within 18 months of infusion. As none of the animals that had received BM transduced with a control lentiviral vector developed leukaemia, it was concluded the γc transgene was the main cause of leukaemia development in these mice. However, in a number of other studies in which γc was expressed at more physiological levels, as had been the case in the X-SCID gene therapy trial, oncogenicity of the γc transgene could not be demonstrated (Thrasher *et al.* 2006). A mouse model of X-SCID gene therapy has now been developed to assess risk factors specific to this disease (See section 1.5.1.)

In a recent series of follow-up papers on the SCID trials, the integration site distribution in pre-transplant $CD34^+$ cells was compared to that in peripheral blood cells of different lineages at varying times post-transplantation (Aiuti *et al.* 2007; Deichmann *et al.* 2007; Schwarzwaelder *et al.* 2007). In all 3 trials, there was an increased frequency of RIS near promoters and CIS in post-transplant samples. This suggests that selective pressure after transplantation allows such clones to engraft better. Identical RIS were found in both lymphoid and myeloid cells of the same patients, indicating HSC were transduced, despite their sparsity in $CD34^+$ populations. RIS near known proto-oncogenes including LMO2 were recovered from healthy patients in whom no clonal imbalance was evident, this is consistent with the idea that several genetic changes are needed for malignant transformation to occur.

1.4.4.2. Leukaemia and clonal dominance in animal models of gene therapy

Several months prior to the report of leukaemias in the French X-SCID trial, the first case of leukaemia in an animal model of gene therapy using a retroviral vector was reported (Li *et al.* 2002). In a murine retroviral gene marking study, acute myeloid leukaemia (AML) was observed in secondary and tertiary

recipients of pooled BM transduced with a retroviral vector expressing a modified Low Affinity Nerve Growth Factor Receptor (dLNGFR). All affected mice carried the same leukaemic clone in which a single copy of the vector had integrated in the first intron of the Evi1 gene in the same transcriptional orientation. Evi1 encodes a zinc finger transcription factor and was first identified as a CIS in myeloid tumours of AKXD23 mice (Mucenski *et al.* 1988). In leukaemic clones, Evi1 was expressed; transcripts were initiated from both LTRs.

In a subsequent study, it was shown leukaemia could also develop in primary recipients of bone marrow, transduced at high MOI with an MDR expressing retroviral vector (Modlich *et al.* 2005). Leukaemia development is thus a function of vector dose. Leukaemic clones contained multiple RIS, at least 3 of which were in putative proto-oncogenes.

In subsequent serial transplantation studies in C57Bl6 mice of retroviral vector transduced bone marrow, no leukaemias developed (Kustikova *et al.* 2005). However, haematopoiesis in secondary recipients relied on just a few clones. RIS in these dominant clones were uniquely in known CIS, proto-oncogenes or other signalling genes, including two in Evi1. Three independent hits in this locus were detected amongst the 22 RIS cloned from primary recipients, but surprisingly none of these Evi1 clones became dominant in secondary recipients.

The first report of malignancy in a large animal model of gene therapy emerged in 2006 (Seggewiss *et al.* 2006). A rhesus macaque developed AML 5 years after gene therapy. Autologous $CD34^+$ cells were transduced with an MSCV vector encoding GFP and a drug resistant variant of the dihydrofolate reductase gene (L22Y). During the first year after transplantation, this animal experienced 80% myeloid marking; most marked cells derived from a single clone. This clone disappeared and for the next 3 years 1-4% of PB cells were GFP positive. In vivo selection for vector transduced cells was performed 2 years after gene therapy. This resulted in a transient increase in marked cells. One of the two vector integrations in the leukaemic clone was found to have occurred in the Bcl2-A1 gene. Interestingly, this very same clone had been responsible for the transient clonal dominance seen in the first year after transplantation.

Administration of SIN EIAV but not SIN HIV vectors into foetal and neonatal mice is associated with the development of liver tumours (Themis *et al.* 2005). Analysis of vector integration sites in these tumours did not reveal targeting of one specific locus, so it is still not known whether insertional mutagenesis played a role in tumour development. EIAV but not HIV vectors in this study contained a longer version of the WPRE element that encoded a truncated X protein. This protein is known to be involved in the development of hepatocellular carcinoma (HCC) in woodchucks. (The WPRE in most vectors, including those in use in our lab, contain mutations that prevent X protein expression.)

In a recent report, HCCs in mice were observed after in vivo neonatal gene therapy with an adeno-associated virus (AAV) vector (Donsante *et al.* 2007). These two studies taken together suggest administration of vectors to the foetal and neonatal liver might be particularly tumourigenic. As can be seen in Table 1.3., AAV infrequently integrates into the genome. The vector used here encoded the β-glucuronidase gene, defective in mucopolysaccharidosis VII, which was expressed from a β-actin promoter and CMV enhancer. AAV vector integration sites were recovered from the tumours of 4 mice. All four were found to have occurred in a 6kb locus on mouse chromosome 12. It now seems that insertional mutagenesis is not restricted to gene therapy with retroviral vectors. Microarray analysis of neighbouring cellular gene expression revealed significant up-regulation of the Rian and Mirg genes, which encode multiple small nucleolar RNAs (snoRNAs) and miRNAs, respectively. Only miRNAs have thus far been implicated in cancer. Insertional activation of the mir-17-92 cluster, a known oncogenic miRNA cistron (primary transcript), was recently found in SL3-3 MLV-induced T cell lymphomas (Wang *et al.* 2006). Insertional activation of this same miRNA cluster was also implicated in some F-MLV induced erythroleukaemias (Cui *et al.* 2007).

1.5. Assays to quantify insertional mutagenesis by retroviral and lentiviral vectors

Despite the large number of animal studies that had preceded the clinical gene therapy trials, the number of leukaemias encountered in the recipients of γc-transduced HSC was not anticipated. In hindsight, the duration of follow-up in these preclinical studies was arguably not long enough and also the levels of gene transfer were too low to observe adverse effects. For the benefit of future gene therapy clinical trials, it is now time for a thorough re-evaluation of the mutagenic potential of vectors. Toxicology studies of any pharmaceutical compound -gene therapy vectors can be classed as such- require suitable assay systems. Like any other assay, those that measure insertional mutagenesis by integrating vectors need to be sensitive, quantitative and reproducible. In general, this means there needs to be a measurable effect of the vector above the background spontaneous mutation frequency; this effect would be expected to be dependent on vector dose, i.e. the number of integrants being screened. Ideally, a direct comparison between vectors based on different parental viruses and vectors of different configuration could be made. Such assays would then allow us to design vectors with a reduced potential for insertional mutagenesis. Although in vivo studies usually follow in vitro studies, animal models developed to assess the potential for insertional mutagenesis by retro- and lentiviral vectors will be discussed first.

1.5.1. In vivo mutagenesis assays

Neonatal wild type Mo-MLV infection can accelerate tumour formation in the Cdkn2a$^{-/-}$ mouse model (Lund *et al.* 2002). This mouse is deficient in p16^{Ink4a} and p19^{Arf}, which regulate the Rb and p53 pathways, respectively. Bone marrow from the Cdkn2a$^{-/-}$ mouse was transduced with either a SIN lentiviral vector expressing GFP from an internal PGK promoter (LV) or a retroviral vector with intact LTRs (RV) (Montini *et al.* 2006). These cells were transplanted into lethally irradiated wild type FVB mice. Twice as many LV as RV integrants were screened, yet in the group that had received LV transduced cells, survival was comparable to that

of control animals. Accelerated tumourigenesis was only observed in the group that had received the highest dose of RV vector. These mice had an increased risk of developing myeloid, but not lymphoid, tumours, which was explained by preferential retroviral LTR activity in myeloid cells. Tumours contained an average of 2-3 copies of the vector per genome.

An X-SCID disease-specific tumour prone mouse model was used to investigate any risk factors unique to X-SCID gene therapy (Shou *et al.* 2006).

In this study, $Arf^{-/-}$ $\gamma c^{-/-}$ double knockout mice were used as BM donors. Total BM cells were transduced with MSCV-γc-IRES-GFP or MSCV-GFP only vectors and then transplanted into lethally irradiated CD45.1 mice. 85% of mice that received γc-IRES-GFP transduced BM developed tumours by 55 weeks of follow-up versus 21% in the control group. The incidence of leukaemia in the γc-IRES-GFP group was much lower when $Arf^{-/-}$ $\gamma c^{+/+}$ BM was used. The following hypotheses were proposed that could account for these differences: There are no NK cells present in $\gamma c^{-/-}$ mice, which are involved in tumour surveillance; IL-7 signalling confers a strong proliferative signal to T cell precursors and thirdly there is an expanded pool of primitive (lymphoid) progenitors in X-SCID.

The following important observations were made in these two studies: In the first study a direct comparison between a non-SIN retroviral vector and a SIN lentiviral vector containing a cellular internal promoter was made. The latter combination was shown to have a reduced potential for insertional mutagenesis, suggesting it might be safer to use in future gene therapy clinical trials. The second study emphasised the importance of the disease background in gene therapy, with X-SCID being a potentially more risky setting than other conditions.

1.5.2. In vitro mutagenesis assays

Prior to tests in animals, the mutagenicity of all potential new pharmaceutical compounds are first tested in in vitro mutagenesis assays. These are designed to be simple, quick and cheap. They will never supersede tests in animals, but provide a convenient first indication of the genotoxic potential of new compounds. In this section, firstly the principles of the simplest prokaryotic mutagenesis systems will be introduced. These measure gain of function following exposure to mutagens. Assays using cultured mammalian cell lines have been in existence for several decades. These measure inactivation of a defined allele following mutagen exposure. As we shall see, retroviruses and derived vectors were tested in these assays.

At present, one in vitro assay, which uses primary cells, has been developed to measure insertional mutagenesis by retro- and lentiviral vectors. The findings of this study will be discussed. Lastly, I will end this introduction by setting out the aims of this thesis.

1.5.2.1. Prokaryotic in vivo mutagenesis assays

The simplest assay to measure the mutagenicity of chemical compounds is the Ames Test (Reviewed in: (Mortelmans and Zeiger 2000)). It was developed in the 1970s by Bruce Ames (Ames *et al.* 1972). The Ames test uses several strains of Salmonella typhimurium that are unable to synthesise histidine. Each strain carries different mutations in the His gene. The idea is that exposure to a mutagen can introduce mutations into the His gene that restores its function. Different mutagens cause different types of mutations, which are picked up by the particular strain in which most revertants are observed. In this way, the Ames test provides limited information about the mechanisms by which a particular compound induces mutations. The number of colonies seen on an agar plate containing minimal histidine is proportional to the mutagenicity of the compound tested. The Ames test has been modified over the years to improve its sensitivity. Indirect toxicity testing was enabled by the inclusion of a metabolic activation system in the test. This is most commonly rat liver homogenate, which provides

enzymes e.g. cytochrome P450, which metabolise compounds to intermediates. This reaction occurs in the presence of the tester strains, so that mutagenicity of the metabolites is assayed. Furthermore, the lipopolysaccharide coat of the tester strain is engineered so it is more permeable to chemicals. Error prone DNA repair is encouraged in some strains, either by deletion of the gene enabling accurate excision repair or by introduction of a plasmid encoding this function. Despite its relative simplicity, the Ames test is still widely used as an initial screen to test the genotoxicity of new compounds.

1.5.2.2. In vivo mutagenesis assays using cultured mammalian cells

Instead of screening for gain of gene function, mutagenesis assays using cultured mammalian cells select for loss of gene function following exposure to a mutagen. The principle is otherwise similar to that of the Ames test. Cells are incubated in the presence of a mutagen. The mutagen can disrupt expression of a defined allele, commonly Hprt or Tk (See below); the absence of the protein encoded by this gene is selected for. Following exposure to the mutagen, cells are firstly cultured for a few days in normal growth medium to allow phenotypic expression of potential mutants; mutants are then selected in the presence of a compound that kills only parental cells.

Just like chemical compounds and physical agents, the mutagenicity of retroviruses and later retroviral vectors were tested in these systems.

The first demonstration that a chronically transforming retrovirus, Mo-MLV, could inactivate a defined allele came in 1981 (Varmus *et al.* 1981). A rat fibroblast cell line that had been transformed with RSV, which was present as a single copy per genome, was superinfected with Mo-MLV. Morphological revertants, that had lost v-src expression, were selected for. In only approximately 10% of revertants, rearranged RSV proviruses could be found. In two revertants, Mo-MLV had integrated into the RSV provirus, upstream of v-src but downstream of the 5'LTR driving its expression.

Mo-MLV was subsequently tested in two well established mammalian mutagenesis systems. In a 1985 study, the mouse F9 embryonal carcinoma cell line was used (King *et al.* 1985). In this system, loss of hypoxanthine-guanine

phosphoribosyltransferase (Hprt) following Mo-MLV infection was selected for. This enzyme is involved in purine salvage. It is encoded by an X-linked gene of ~33kb. In male cells it is thus present as a haploid allele. Hprt $^{-/-}$ cells can be grown in medium containing the purine analogue 6-thioguanine (6TG). Hprt $^{+/-}$ cells incorporate this compound into DNA, where it leads to strand breaks and cell death. Following Mo-MLV infection, Hprt $^{-/-}$ mutants were obtained at a cell frequency 12 fold over the background frequency of 3.5 x 10^{-8}. The provirus frequency was 1-4 x 10^{-8}. More mutants were obtained at higher virus copy numbers.

A more recent study measured Hprt loss after retroviral vector transduction of V79 Chinese hamster cells (Themis *et al.* 2003). The spontaneous frequency to Hprt $^{-/-}$ in these cells is 2.9 x10^{-6}. MLV vector transduction at low MOI of 1-2, did not increase the mutation frequency over background; transduction at high MOI (~6), did increase the mutation frequency 2.3 fold over background. It is not clear how many vector integrants were screened in this study. 63% of Hprt $^{-/-}$ clones analysed had rearranged the Hprt allele.

MLV based vectors were also used to measure thymidine kinase (TK) inactivation in the TK6 human B cell lymphoma cell line (Grosovsky *et al.* 1993). Selection of TK $^{-/-}$ mutants was performed in the presence of trifluorothymidine (TFT), a pyrimidine analogue, that is incorporated into DNA of TK $^{+/-}$ cells. Vector transduction at low MOI of 1-2, resulted in a 5 fold increase in the cell frequency at which mutants were obtained. Again, in only a minority of mutants could an integrated provirus in the TK locus be detected.

As was demonstrated in section 1.2.2.4., mutagenicity associated with retroviruses and vectors derived from them, is not just confined to the disruption and inactivation of single copy genes. The types of mammalian cell culture assays described above did therefore underestimate the mutagenicity of retroviruses and retroviral vectors. As shall become clear in the introduction to Chapter 3, the assay we are proposing to develop has more in common with the Ames test in the sense that it selects for gain of function mutations. Gain of

function of a range of different target genes by retroviral vector insertion is selected for.

1.5.2.3. Immortalisation of primary bone marrow by retroviral vectors

It was demonstrated that immortalised cell lines could be obtained after serial passage of primary BM cells that had been transduced with a murine stem cell virus (MSCV) - based vector (Du *et al.* 2005). These cell lines were immature myeloid cells in phenotype; they remained dependent on growth factors and were unable to induce leukaemia when transplanted back into mice. Vector integration sites in immortalised clones were identified. Half contained RIS in Evi1 or Prdm16, resulting in upregulated expression of these genes.

The Baum lab adapted this approach to develop an in vitro assay for insertional mutagenesis (Modlich *et al.* 2006). Lineage negative primary BM cells from C57Bl6 were transduced; then grown as a bulk culture for 2 weeks, before being plated out into 96 well plates. The number of wells containing clones that had acquired serial replating ability was scored. Two different vectors were initially compared: a retroviral vector containing wild type 5' and 3' SFFV LTRs (WT SFFV) and a SIN retroviral vector containing an internal SFFV U3 promoter (SIN SFFV). The replating frequency per vector copy number was 12 fold lower for the SIN compared to the wild type retroviral vector (1 in 362 versus 1 in 4347 transducing events). In a follow-up study, a SIN vector containing an internal Elongation Factor 1α (EF1α) promoter (SIN EFS) was tested (Zychlinski *et al.* 2008). No clones were detected with this vector; the replating frequency of this vector was estimated to be at least 10 fold lower than that of SIN SFFV. A SIN vector containing a 250bp chicken cHS4 insulator (SIN SFFV cHS4) in the U3 region was also tested. The replication frequency per vector copy of SIN SFFV cHS4 was 2.5 fold lower than that of SIN SFFV. The insulator element thus somewhat reduces insertional mutagenesis, but unlike the SIN EFS vector there is still detectable transforming activity.

Clones obtained with the wild type vector shared many RIS. In 8 clones analysed in detail, 6 contained independent RIS near Evi1; all 8 had upregulated Evi1 expression. 2 out of 3 SIN-SFFV clones analysed, also had activating insertions

in Evi1. This cell culture assay suggests that an internal EF1α promoter is unable to transactivate Evi1 expression. This gives some indication that SIN vectors containing cellular promoters might be safer to use in gene therapy.

1.6. Thesis Aims

This thesis describes the development of a cell line assay that allows us to quantify insertional mutagenesis by retroviral and lentiviral vectors and to analyse the mechanisms by which this occurs.

The interleukin-3 (IL-3) dependent cell line BCL15 is derived from the mouse bone marrow cell line BAF3 and over-expresses human Bcl2. The frequency at which IL-3 independent mutants are obtained following vector transduction was measured.

I aimed to compare the mutagenic potential initially of lentiviral versus retroviral vectors and subsequently of different vector configurations. As will be shown in chapter 3, insertional activation of one of several cellular genes can transform BAF3 cells to IL-3 independence. Which of these genes had been targeted by vector insertion in transformed cells was investigated. The mechanisms whereby vector insertion activates expression of transforming genes was analysed.
This in vitro assay may inform us which viral vectors and vector configurations are safest. This will allow the design of vectors with a reduced mutagenic potential for use in future gene therapy pre-clinical and clinical trials.

Chapter 2

2. Materials and methods

All buffers and solutions named in the text are listed in Table 2.1.

2.1. Tissue culture materials and methods

2.1.1. Tissue culture adapted cell lines

The following human and mouse derived cell lines were used: 293T, TE671, BAF3 (Ba/F3), BCL15 and WEHI3B.

The 293T cell line is a highly transfectable derivative of the 293 human embryonic kidney cell line, that expresses the SV40 large T antigen. The TE671 cell line is derived from a human rhabdomyosarcoma. They were maintained in Dulbecco's Modified Eagle's Medium (DMEM) (Gibco), supplemented with 10% Foetal Calf Serum (FCS), 2mM L-Glutamine, 100 units/ml Penicillin and 100 µg/ml Streptomycin (all from Gibco). This medium will be referred to as complete medium. The 293T and TE671 cell lines are both adherent; cells were subcultured using trypsin/EDTA (Gibco). 293T cells were passaged 1:4 or 1:6 every 2-3 days; TE671 were passaged 1:10 or 1:20 every 3-4 days.

WEHI3B is a BALB/c mouse myelomonocytic leukaemia cell line that produces IL-3 as a result of insertional activation of the mouse IL-3 gene by an intracisternal A particle (IAP) (Ymer *et al.* 1985). Conditioned medium was harvested from WEHI3B cells when the cells reached confluence (cell density of 10^6 cells/ml). WEHI3B supernatant was filtered through a 0.45 µm filter and stored in aliquots at -20°C. The optimal concentration of WEHI3B supernatant needed to grow BAF3 cells was determined by tritium [3H] thymidine proliferation assay.

The BAF3 cell line is an IL-3-dependent mouse pro-B lymphocyte cell line (Palacios and Steinmetz 1985). The BCL15 cell line, which is also dependent on IL-3, is derived from BAF3, but already over-expresses human Bcl2 (Collins *et al.* 1992). BAF3 and BCL15 cells grow in suspension; both cell lines were grown in complete medium supplemented with 2-10% WEHI3B cell conditioned medium as a source of IL-3. BAF3/BCL15 cells were passaged 1:20 every 2 days or 1:50 every 3 days, to maintain cell density between 5 x 10^4 and 10^6 cells/ml. All cell lines were grown in a 10% CO_2 incubator at 37°C.

2.1.2.Production of VSV-G pseudotyped retroviral and lentiviral vectors

2.1.2.1. Transfer vector plasmids

The HIV-1-derived lentiviral vectors used were pHV (Ikeda *et al.* 2003) and pHRSIN-CSGW (Demaison *et al.* 2002) (Figure 2.1.a.). pHV has a wild-type 3' long terminal repeat (LTR) from HIV-1 subtype B isolate HXB2. pHV's 3' and 5'LTRs were derived from pH7G (Oxford Biomedica). Its self-inactivating derivative pHRSIN-CSGW, has a 400 bp deletion in the 3'LTR U3 region. These vectors are otherwise identical and contain an internal spleen focus-forming virus (SFFV) LTR promoter driving eGFP expression and a woodchuck hepatitis virus post-transcriptional regulatory element (WPRE). pHRSIN-CSGW was kindly provided by A. Thrasher (Institute of Child Health, UCL).

The retroviral vector pCNCG (Soneoka *et al.* 1995) contains wild type Mo-MLV LTRs and an internal CMV promoter driving GFP expression. The MPSV LTR-driven retroviral vector pMFG.S eGFP (Riviere *et al.* 1995) was kindly provided by J.S. Lee (Harvard Gene Therapy Initiative) (Figure 2.1.b.).

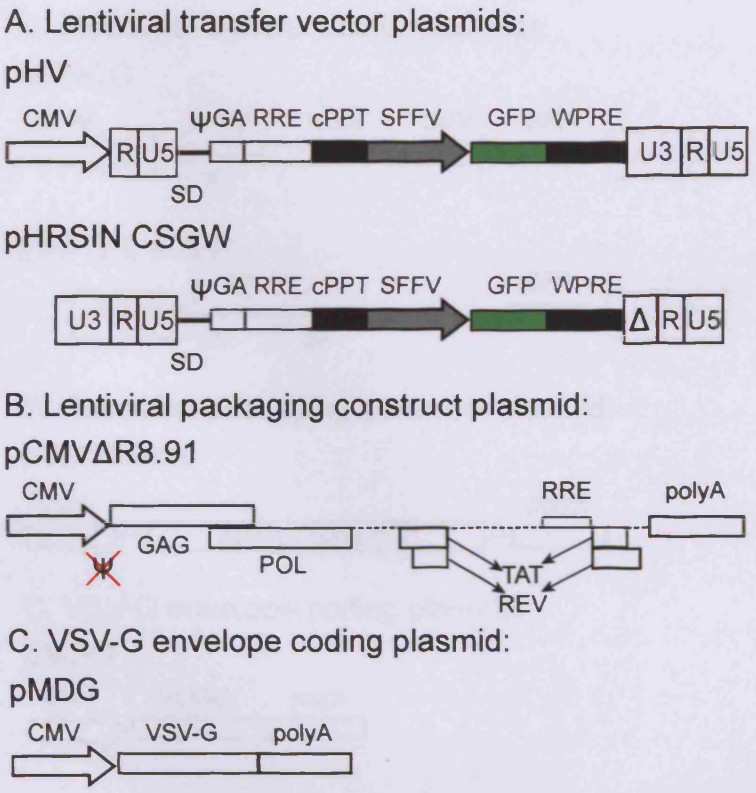


Figure 2.1.a. Constructs used in transient three plasmid transfection of 293T cells for the production of high titre lentiviral vector.

A. The HIV-1 derived lentiviral transfer vectors pHV (Ikeda et al, 2003) and pHRSIN CSGW (Demaison et al, 2002) are shown. pHV was a wild type HIV-1 3'LTR, whereas pHRSIN CSGW contains a 400bp deletion (Δ) in the U3 region of the HIV-1 3'LTR, resulting in the generation of a self-inactivating lentiviral vector. The vector backbones of pHV and pHRSIN CSGW are otherwise identical.
Lentiviral transfer vector plasmids were co-transfected into 293T cells with pCMVΔR8.91 (B) and pMDG (C).
B. pCMVΔR8.91 (Zufferey et al, 1997) is a second generation lentiviral packaging construct expressing HIV-1 Gag, Pol, Tat and Rev.
C. The envelope coding plasmid pMDG (Naldini et al, 1996) is used to pseudotype lentiviral vector particles with the vesicular stomatitis virus G (VSV-G) protein.

CMV is the human cytomegalovirus immediate early promoter; Ψ is lentiviral packaging signal; SD is HIV-1 major splice donor; GA is truncated Gag 5'UTR; RRE is Rev-responsive element; cPPT is central polypurine tract; SFFV is spleen focus forming virus LTR promoter/enhancer; GFP is green fluorescent protein reporter gene; WPRE is woodchuck post-transcriptional regulatory element.

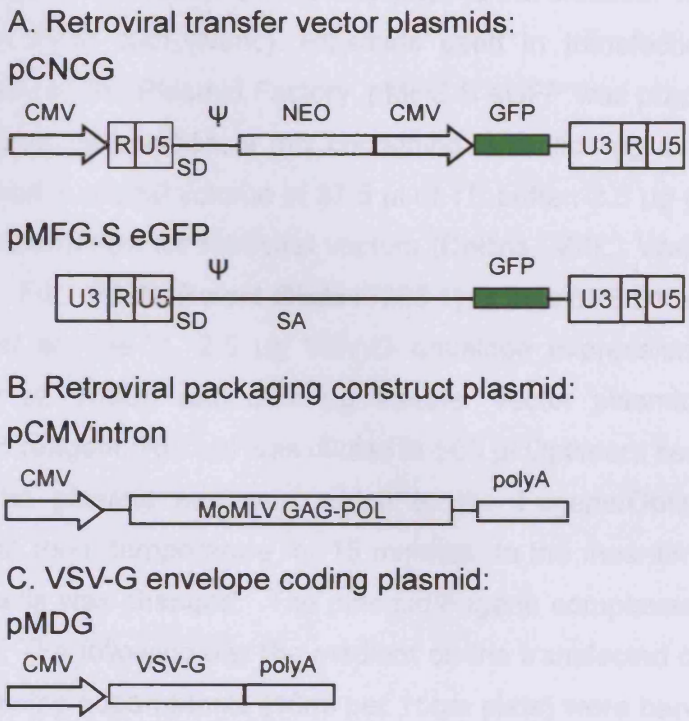


Figure 2.1.b. Constructs used in transient three plasmid transfection of 293T cells for the production of high titre retroviral vector.

A. The retroviral transfer vectors pCNCG (Soneoka et al, 1995) and pMFG.S eGFP (Riviere et al, 1995) are shown. These constructs contain wild-type Moloney Murine Leukaemia Virus (Mo-MLV) and myeloproliferative sarcoma virus (MPSV) 3'LTRs, respectively.
Retroviral transfer vector plasmids were co-transfected into 293T cells with pCMVintron (B) and pMDG (C).
B. pCMV intron encodes Mo-MLV Gag-Pol.
C. pMDG was used to produce VSV-G pseudotyped retroviral vector particles.

CMV is the human cytomegalovirus immediate early promoter; Ψ is retroviral packaging signal; SD is splice donor; SA is splice acceptor; NEO is the neomycin resistance gene; GFP is the green fluorescent protein reporter gene.

2.1.2.2. Three plasmid transfection of 293T cells

Vector virus was produced by transient three-plasmid transfection (Soneoka *et al.* 1995; Besnier *et al.* 2002). On the day prior to transfection, 10^7 293T cells were plated per 15cm dish (Nunc). Plasmids used in transfection were prepared commercially by the Plasmid Factory. pMFG.S eGFP was prepared by Maxiprep. On the day of transfection, a mix containing the following quantities of plasmid was prepared in a total volume of 37.5 µl of TE buffer: 2.5 µg gag-pol expression plasmid (pCMVintron for retroviral vectors (Collins, MKL., Weiss, RA., Takeuchi, Y., Cosset, F-L. 1995, Patent GB9517263.1); pCMVΔR8.91 for lentiviral vectors (Zufferey *et al.* 1997)), 2.5 µg VSV-G envelope expression plasmid (pMD-G (Naldini *et al.* 1996)) and 3.75 µg transfer vector plasmid. 45 µl Fugene6 transfection reagent (Roche) was diluted in 500 µl Optimem serum-free medium (Gibco). The plasmid mix was added to the Fugene/Optimem mixture and incubated at room temperature for 15 minutes. In the meantime, the medium on the 293T cells was changed. The plasmid/Fugene complexes were then added to the cells. The following day the medium on the transfected cells was changed. Virus containing supernatants (18ml per 15cm plate) were harvested at 48h, 72h and 96h post-transfection, filtered through 0.45µm filters and either concentrated immediately or stored at - 80°C. Vector virus was 100 fold volume concentrated by ultracentrifugation for 2h at 19.000 rpm at 4°C in a Sorvall ultracentrifuge using the Surespin (6 x 36ml) swinging bucket rotor. Virus was resuspended in cold complete medium and left on ice for 2h before aliquoting and storage at - 80°C.

2.1.2.3. Titration of retroviral and lentiviral vector stocks on 293T and BAF3/BCL15 cells

Each batch of vector virus was titered on TE671 and BAF3/BCL15 cells. TE671 cells were plated onto six-well culture dishes at a density of 10^5 cells/well one day prior to transduction. The number of cells in one well was counted just before transduction and this cell number was used in titre calculations. Serial dilutions of vector virus were prepared in DMEM supplemented with 5 µg/ml Polybrene (PB, hexadimethrine bromide, Sigma). 1 ml of each dilution was then added to each well of TE671 cells. The medium on the cells was changed the following day. For titrations on BAF3/BCL15 cells, serial dilutions of vector virus were once again prepared. 5 x 10^5 Baf3 cells were infected in a total volume of 500 µl medium in the presence of 5µg/ml PB. BAF3/BCL15 cells were incubated in the presence of virus for 4h. After 4h, the culture was expanded to a total volume of 10ml. Both TE671 and BAF3/BCL15 cells were assayed for GFP expression by flow cytometry at 72h post-transduction. Virus input was plotted against the percentage of GFP positive cells (See Figure 2.2. for a typical titration curve). The titre in infectious units (i.u.) per ml was calculated from data points where approximately 10% of the cells were GFP positive (Table 2.2.). Lentiviral and retroviral vector titers were in the range of 1 - 5 x 10^7 and 5 - 10 x 10^6 BCL15 (GFP) infectious units/ml, respectively. Some vector batches were also titered by Q-PCR for integrated GFP copies. 10^6 cells were harvested at 72 hours post-transduction and DNA was extracted using a DNAeasy kit (Qiagen).

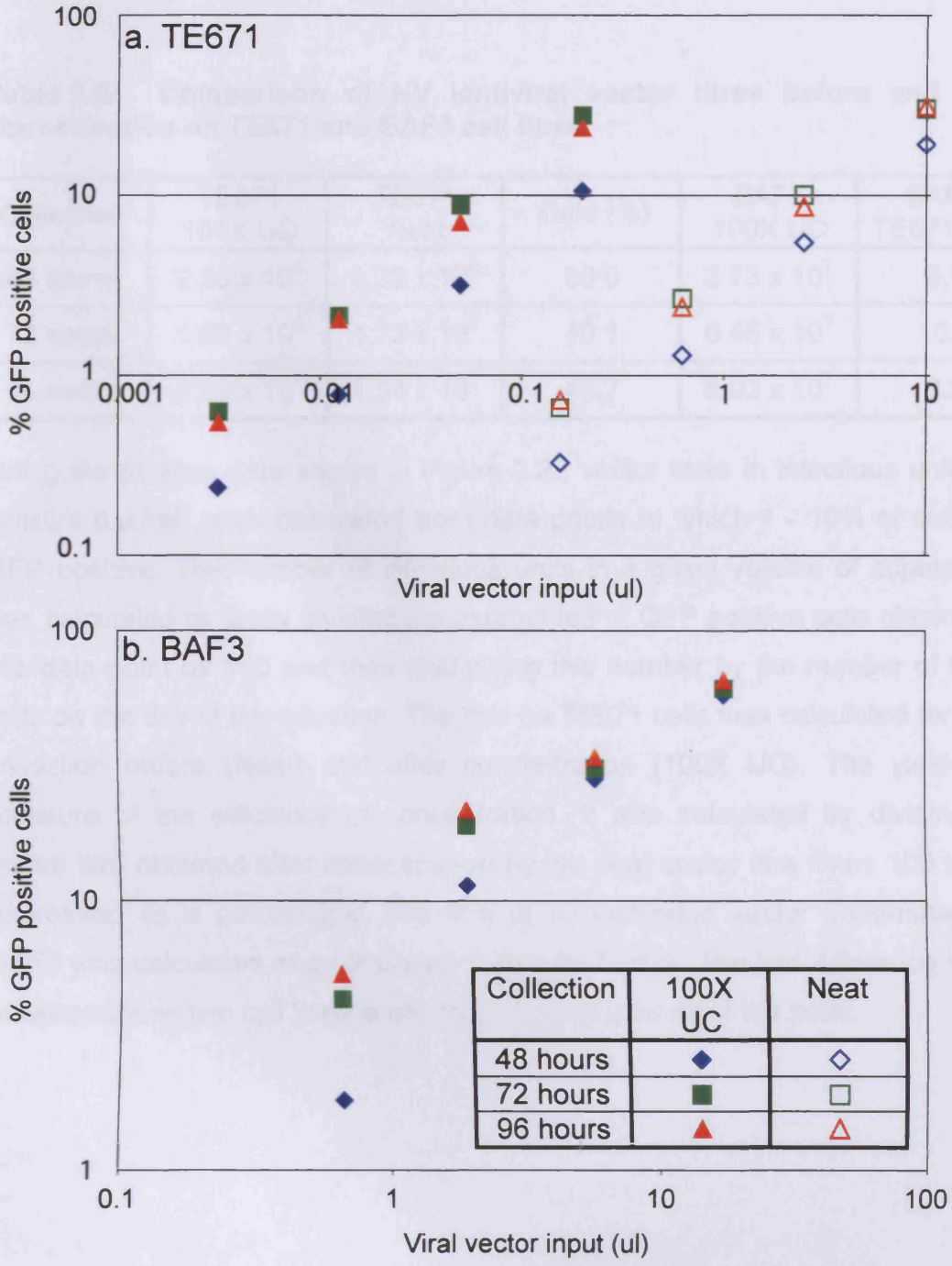


Figure 2.2. Titrations of neat and 100 fold volume concentrated HV vector supernatant on TE671 and BAF3 cells

HV vector containing supernatant was collected at 48, 72 and 96 hours post-transfection. Four fold serial dilutions of neat supernatant (open symbols) or supernatant 100 fold volume concentrated by ultracentrifugation (filled symbols) were added to 4 x 10^5 TE671 cells (A) or 10^6 BAF3 cells (B, concentrated supernatant only). Cells were transduced in the presence of 5μg/ml Polybrene (hexadimethrine bromide). The percentage of GFP positive cells was measured by flow cytometry at 48 hours post-transduction. Vector titres were calculated from data points at which 1 – 10% of cells express GFP and are shown in Table 2.2.

Table 2.2. Comparison of HV lentiviral vector titres before and after concentration on TE671 and BAF3 cell lines

Collection	TE671 100X UC	TE671 Neat	Yield (%)	BAF3 100X UC	BAF3/ TE671 ratio
48 hours	2.50×10^8	8.32×10^6	30.0	2.73×10^7	9.1
72 hours	6.88×10^8	1.72×10^7	40.1	6.45×10^7	10.7
96 hours	6.59×10^8	1.54×10^7	42.7	8.03×10^7	8.2

Using the titration data shown in Figure 2.2., vector titres in infectious units per millilitre (i.u./ml) were calculated from data points at which 1 - 10% of cells are GFP positive. The number of infectious units in a given volume of supernatant was calculated by firstly dividing the percentage of GFP positive cells obtained at this data point by 100 and then multiplying this number by the number of target cells on the day of transduction. The titre on TE671 cells was calculated for each collection before (Neat) and after concentration (100X UC). The yield is a measure of the efficiency of concentration. It was calculated by dividing the vector titre obtained after concentration by the neat vector titre times 100 and is expressed as a percentage. The titre of concentrated vector supernatant on BAF3 was calculated and compared to that on TE671. The fold difference in titre between these two cell lines is shown in the last column of the table.

2.1.3. Irradiation experiments

5 x 10^6 parental BAF3 and HV A2 cells were washed thrice in complete medium without IL-3. Cells were then irradiated using a Pantac X-ray machine with an output of 240 KeV at a dose rate of 0.7 Gy/min. Immediately following irradiation, cells were pelleted and resuspended in the appropriate volume of medium with or without IL-3. BAF3 cells were cultured at 2 x 10^5 cells/ml (with IL-3) or 10^6 cells/ml (without IL-3). HV A2 cells were cultured at 2 x 10^5 cells/ml (with or without IL-3).

2.1.4. Spiking experiments

BAF3 and HV A2 cells were washed thrice in complete medium without IL-3. Cells were counted and 10 fold serial dilutions of HV A2 cells were prepared. The required number of HV A2 cells (in 500µl) was added to 50ml culture containing 5 x 10^7 BAF3 cells. In the experiment shown in Figure 6.3., spiked cultures were irradiated at 2 Gy using the Pantac X-ray machine as above.

2.1.5. Cell viability measurements

2.1.5.1. Trypan blue exclusion

One volume of cell culture was incubated for 5 minutes with an equal volume of trypan blue. The number of viable, trypan blue excluding, cells was counted in a haemocytometer under a light microscope.

2.1.5.2. Propidium iodide exclusion

Cell viability was measured by propidium iodide (PI) exclusion as described (Thomas *et al.* 2004). 10^5 BAF3 cells were pelleted and then incubated for 5 minutes with HBBS containing 5 µg/ml PI. Uptake of the dye was measured by flow cytometry using a BD LSR and Cell Quest software. A FSC/SSC gate was used to exclude debris. 10.000 cells were analysed on a FSC/FL-2 dot plot. Live cells were FL-2 negative and FSC high; apoptotic cells were FL-2 dull, FSC intermediate and dead cells were FL-2 bright, FSC low. To differentiate between live parental BAF3 cells (GFP negative) and mutants (GFP positive), cells were analysed on a FL-1/FL-2 dot plot in the spiking and mutagenesis experiments.

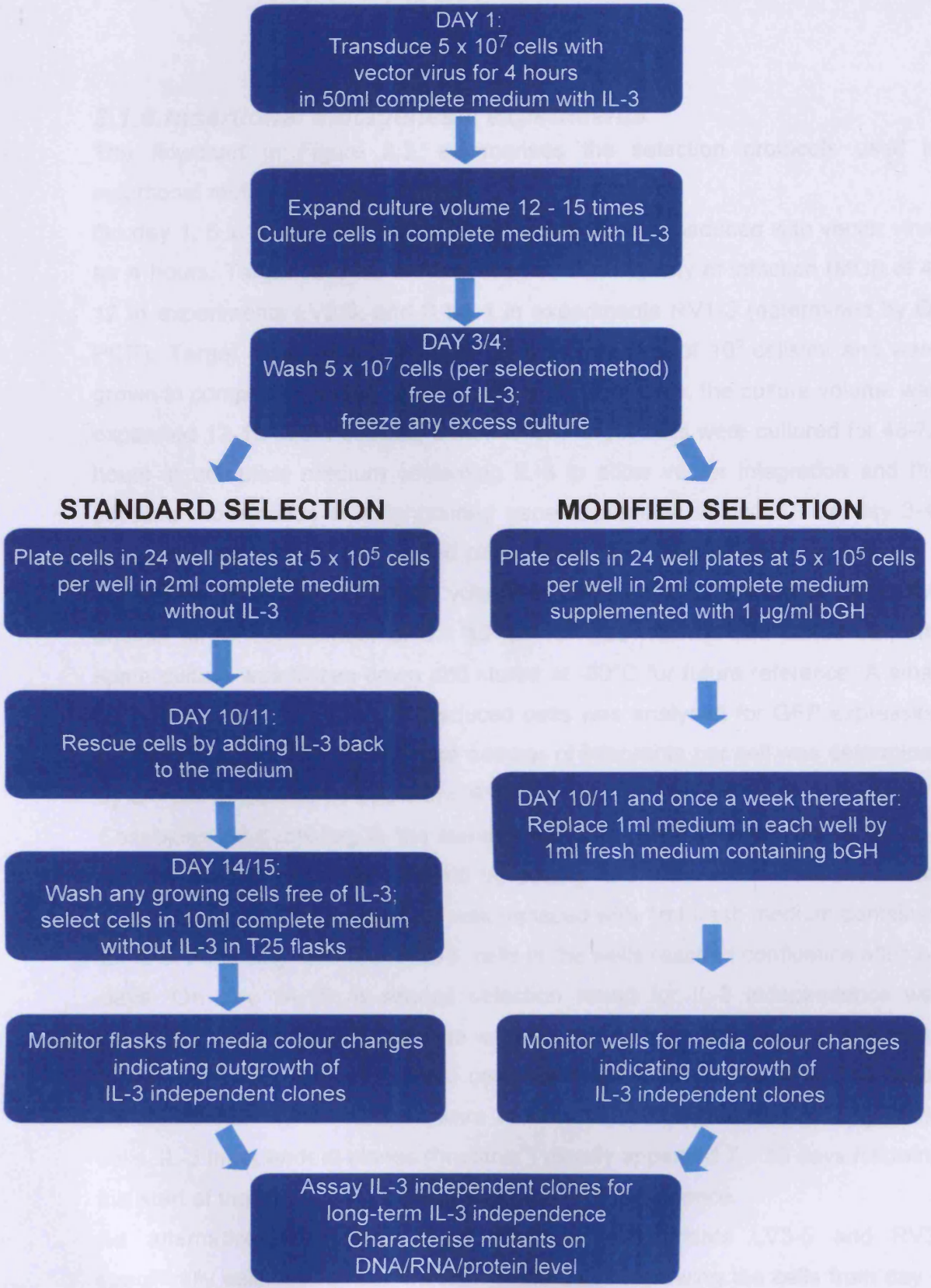


Figure 2.3. Selection protocols used in insertional mutagenesis experiments

2.1.6. Insertional mutagenesis experiments

The flowchart in Figure 2.3. summarises the selection protocols used in insertional mutagenesis experiments.

On day 1, 5 x 10^7 to 10^8 BAF3 or BCL15 cells were transduced with vector virus for 4 hours. Target cells were transduced at a multiplicity of infection (MOI) of 4-12 in experiments LV2-5; and 0.1-1.1 in experiments RV1-3 (determined by Q-PCR). Target cells were transduced at a cell density of 10^6 cells/ml and were grown in complete medium containing IL-3. After 4 hours, the culture volume was expanded 12-15 fold. Following transduction, target cells were cultured for 48-72 hours in complete medium containing IL-3 to allow vector integration and the possible modification of neighbouring gene expression to occur. On day 3-4, cells were washed free of IL-3 and plated in 24 well plates at a cell density of 5 x 10^5 cells per well in 2ml culture volume in complete medium without IL-3. The original target cell number of 5 x 10^7 to 10^8 was washed and plated out; any spare culture was frozen down and stored at -80°C for future reference. A small fraction of mock and vector transduced cells was analysed for GFP expression by flow cytometry and the average number of integrants per cell was determined by Q-PCR (experiments LV2-5 and RV1-3 only).

Cells selected according to the standard selection protocol were starved of IL-3 for 7 days. Cultures were rescued by adding IL-3 back to the medium on day 10/11. In each well, 1ml medium was replaced with 1ml fresh medium containing 2x IL-3. Following rescue with IL-3, cells in the wells reached confluence after 2-3 days. On day 14-15, a second selection round for IL-3 independence was initiated. Cells from each well were washed thrice in complete medium without IL-3 and were then grown in 10ml complete medium without IL-3 in T25 tissue culture flasks. These cultures were monitored for the presence of outgrowing cells. IL-3 independent clones ("mutants") usually appeared 7 – 10 days following the start of the second selection round for IL-3 independence.

An alternative selection protocol, used in experiments LV3-5 and RV3, specifically selected for Ghr-insertional mutants by growing the cells from day 4 onwards in complete medium without IL-3, but supplemented with 1µg/ml

recombinant bovine growth hormone, bGH (A kind gift from G. Bogosian, Research Director in Agricultural Biotechnology, Monsanto, St. Louis, MO, USA). Cells were washed thrice in complete medium, and then resuspended in complete medium containing bGH. Cells were plated at 5 x 10^5 cells per well in 2ml medium. Plates were regularly inspected for media colour changes in the wells, which indicated the outgrowth of a mutant. Mutants usually appeared 2 weeks after the start of selection in serum plus growth hormone.

Mutants were tested for long-term IL-3 independence and further characterised on DNA/RNA/protein level. "Sibling" clones were eliminated after Southern blot, if their band patterns were similar. Once established, any long term IL-3 independent mutants of the BAF3/BCL15 cell line were routinely cultured in complete medium without IL-3.

2.1.7. Statistical analysis

The Chi Square statistical method was used to determine how much the observed cell frequencies differed from the theoretically expected cell frequencies in the mock and vector transduced arms of each experiment.

X^2 values were calculated using Smith's Statistical Package (version 2.80, September 26, 2005, copyright 1995-2005, Gary Smith). In these comparisons, a X^2 value greater than 11 is equivalent to $p<0.001$, which was considered significant.

2.1.8. Phenotypical studies on IL-3 independent mutants

2.1.8.1. Culturing cells in serum free medium or in the presence of recombinant cytokines

Cells were washed thrice in serum free medium. After the final wash, cells were resuspended in complete medium, complete medium containing IL-3, serum free medium or serum free medium containing recombinant cytokines. Recombinant cytokines used were: mouse interleukin 3 (403-ML/CF, R&D systems), mouse insulin-like growth factor 1 (791-MG, R&D systems) and bovine growth hormone (Monsanto).

2.1.8.2. [3H] thymidine proliferation assay

On day 1, cells were washed thrice in serum-free medium and then cultured in serum free medium for 6 hours to synchronise cells and to remove any remaining IL-3 bound to them. Cells were then plated in 96 well flat bottom plates at a density of 2 x 10^4 cells per well in the presence of different concentrations of recombinant cytokines in 200µl total culture volume. On day 2, cells were pulsed with 1µCi [3H]-thymidine per well (ICN Biomedical, High Wycombe, UK). On day 3, DNA from cells was harvested on a filtermat using a Tomtec Harvester Model 96. [3H] incorporation into DNA was measured by liquid scintillation counting (Microbeta Systems). Control wells contained BAF3/BCL15 cells in serum free medium only, and these yielded <100 counts per minute [3H]-thymidine incorporation.

2.1.9. Flow cytometry cell surface staining

5 x 10^5 cells per stain were washed twice in HBSS 2% FCS. Cells were then incubated on ice for 1 hour with 5µg/ml biotinylated goat anti-mouse GHR antibody (BAF1360, R&D systems) which was diluted in blocking buffer (1X PBS, 2% FCS, 0.01% Na Azide). Biotinylated normal goat IgG (BAF108, R&D systems) was used as isotype control; biotinylated goat anti-mouse IL-3 Rβ antibody (BAF549, R&D systems) was used as positive control. Cells were washed twice in blocking buffer. Cells were then incubated with 1 µg/ml Allophycocyanin (APC)-conjugated streptavidin (17-4317, eBioscience) for 30 minutes on ice in the dark. Finally, the samples were washed twice in blocking buffer before being resuspended in 500 µl HBSS. The samples were analysed by flow cytometry using a BD LSR and Cell Quest software.

2.2. Molecular biology techniques

2.2.1. Plasmid preparation and manipulation

2.2.1.1. Preparation of heat shock competent E.coli

XL1-Blue competent cells (Invitrogen) were streaked onto an LB agar plate containing 10 µg/ml tetracycline and grown overnight at 37°C. A colony was picked and inoculated into 2ml LB broth containing 10 µg/ml tetracycline. This culture was grown overnight in a shaking incubator at 37°C. The following morning, 1ml of overnight culture was inoculated into 100ml LB broth without antibiotics. This was grown in a shaking incubator at 37 °C until the OD600 was 0.3-0.6. The culture was cooled on ice for 5 minutes; then the bacteria were pelleted and re-suspended in 50ml ice-cold TFB-I buffer. After a 5 minute incubation on ice, bacteria were pelleted again and re-suspended in 5 ml ice-cold TFB-II buffer. Bacteria were aliquoted after a 15 minute incubation on ice and stored at -80°C.

2.2.1.2. Introduction of plasmid DNA into E.coli

Competent cells were thawed on ice and added to 2 µl ligation reaction. This mixture was left on ice for 20 minutes. Bacteria were heat shocked for 45 seconds at 42°C and immediately put back on ice for 2 minutes. 950 µl of SOC medium (Invitrogen) was added, and bacteria were grown in a shaking incubator at 37°C for 1.5 hours. 100 µl of transformation mixture was then plated on LB agar plates containing the appropriate antibiotics. Blue white screening for colonies containing insert carrying plasmids was performed by precoating LB agar plates with Blue White screening reagent (X-gal/ IPTG) (Sigma). White colonies were further screened by PCR to confirm presence of the insert.

2.2.1.3. Plasmid DNA mini-/midi-/maxi- preparations

Plasmids were extracted using Qiagen plasmid purification kits (QIAprep Spin Miniprep kit; Plasmid Midi kit; Plasmid Maxi kit). A single colony was inoculated into 2-3 ml LB broth containing 100µg/ml ampicillin (Miniprep). The culture was grown overnight in a shaking incubator (250 rpm/min). For midi and maxi preparations a 2 ml overnight starter culture was diluted into 25 or 100 ml selective LB broth and grown overnight in a shaking incubator. Bacteria were pelleted and re-suspended in 250 µl, 4 ml or 10 ml buffer P1. The cells were then lysed by addition of 250 µl, 4 ml or 10 ml buffer P2 and left at room temperature for 4-5 minutes. The mixture was neutralised by addition of 250 µl, 4 ml or 10 ml buffer N3. The resulting precipitate was then pelleted by centrifugation. The supernatant containing plasmid DNA was applied to an anion-exchange resin column provided by QIAprep Mini, Midi or Maxi kits. Purified plasmid DNA was eluted in water or EB buffer. The plasmid DNA concentration was determined by Nanodrop. The quality of the plasmid preparation was determined by running a fraction out on an agarose gel.

2.2.2. Restriction enzyme digests

Restriction enzymes used were obtained from Promega or New England Biolabs (NEB) and were performed according to the manufacturer's instructions. Digests on plasmid DNA were performed for 2 hours. Genomic DNA was digested overnight; to ensure complete digestion extra enzyme was added to overnight digests the following morning.

2.2.3. Isolation of DNA fragments from agarose gels (gel-purification)

1-2% agarose gels were used to separate DNA fragments according to size. Gels were made by dissolving 1-2 grams agarose in 1X TAE buffer (Tris-acetate-EDTA). 0.2µg/ml ethidium bromide was added to visualise DNA on a UV transluminator. 6X DNA loading buffer was added to DNA fragments (and PCR products). 10µl of 100bp or 1kb GeneRuler DNA ladder (Fermentas) was included on each gel to allow quantification and size determination of DNA

fragments. DNA fragments were cut out of agarose gels and purified using a QIAquick gel extraction kit (Qiagen). Agarose gel slices were solubilised; applied to a column carrying a silica membrane; washed and eluted in water or EB buffer.

2.2.4. Subcloning of DNA fragments

PCR products were cloned into pGEM T Easy vector (Promega). 3 µl PCR product was ligated with 50 ng pGEM T Easy vector for 1 hour at room temperature or overnight at 4°C. 2 µl ligation reaction was used to transform competent XL1-Blue E.coli. M13 forward or reverse primers were used to sequence plasmids carrying inserts.

2.2.5. Genomic DNA extraction

Cells were harvested and washed twice with HBSS. Cell pellets were either stored at -80°C for DNA extraction at a later date or used immediately. Genomic DNA for use in quantitative PCR and integration site PCR was extracted from 10^6 cells per sample using a DNeasy Blood and Tissue kit (Qiagen). Cells were lysed for 10 minutes at 56°C using Proteinase K. RNA was digested by adding RNAse A. The samples were loaded onto a DNeasy spin column, washed twice with PE buffer and finally eluted in 200 µl buffer AE.

DNA for use in Southern blotting was prepared from 2 x 10^7 cells. Cells were re-suspended in 500 µl lysis buffer and lysed overnight in a 60°C water bath. To precipitate DNA, 1250 µl of precipitation buffer was added per sample. This was left at room temperature for 30 minutes. The DNA was then pelleted by centrifugation at 13.000 rpm for 15 minutes at 4°C. This was followed by 3 washes of the DNA pellet in 70% ethanol, with 30 minutes shaking time and a 15 minute spin at 13.000rpm per wash. The DNA pellet was air dried for ~10-15 minutes. The pellet was finally resuspended in 200 µl TE pH 8 buffer and left to dissolve at 4°C overnight.

Genomic DNA samples were stored in aliquots at 4°C or -20°C. DNA concentration was measured by NanoDrop spectrophotometer; a small fraction was also analysed by gel electrophoresis to assess genomic DNA quality.

2.2.6. Southern blotting

15 µg genomic DNA was digested overnight at 37°C with 60 units of a restriction enzyme in a total volume of 200 µl. Digested DNA was concentrated by ethanol precipitation and dissolved in 20 µl TE buffer. DNA fragments were separated by overnight gel electrophoresis on a 0.7% agarose gel at 20V. A picture was taken of the gel, and the transfer and hybridisation protocol was started immediately. Prior to blotting the DNA from the gel onto a membrane, the gel was sequentially submerged in depurination solution for 10 minutes, denaturation solution for 30 minutes and neutralisation solution of 30 minutes. Each of these incubations was done on a revolving platform at room temperature. The gel was rinsed with distilled water in between incubations with the different solutions. DNA fragments were blotted overnight by neutral transfer, using 20X SSC, onto a positively charged nylon membrane (HybondTM–XL, Amersham Bioscience). DNA fragments were crosslinked onto the membrane using a Stratalinker UV Crosslinker (Stratagene). The membrane was then washed in 2X SSC. The membrane was pre-hybridised in Rapid-HybTM Buffer (GE Healthcare) supplemented with 10 µg/ml Salmon Testis DNA (Sigma) overnight at 65°C. Digested genomic DNA was probed with a 32P labelled GFP probe (A 720bp PCR product generated using primers GFP-F and GFP-R; see Table 2.3.1.). The probe was made using the Rediprime II Random Prime Labelling System (GE Healthcare). Each probe was labelled with 50 µCi (=1.85 MBq) of Redivue [α-32P] dCTP (GE Healthcare). Unincorporated nucleotides were removed using the QIAquick Nucleotide Removal kit (Qiagen). Prior to addition to the hybridisation buffer, the probe was denatured by heating it to 100°C for 5 minutes. The membrane was hybridised in the presence of the probe overnight at 65°C. The following day, the membrane was washed thrice in wash solution 1 (1 x for 10 min and 2 x for 15 min) and once in wash solution 2 for 15 min at 65°C. The membrane was exposed to a phosphorscreen for several days. The phosphorscreen was visualised using a Storm PhosphorImager (GE Healthcare). The image was analysed using ImageQuant software (GE Healthcare). Following exposure to phosphorscreen, blots were exposed to Kodak BioMax film (Kodak).

2.2.7. Polymerase chain reaction (PCR)

All polymerase chain reactions were performed using HotStarTaq DNA polymerase (Qiagen). This enzyme has no polymerase activity at room temperature and requires a 15 minute activation step at 95°C. PCR reactions were performed in a total reaction volume of 25 to 100 µl. PCR reaction components and amplification parameters are shown below. PCR reactions were performed using a Peltier thermal cycler (Dyad) equipped with a heated lid. Primers were obtained from Invitrogen and are listed separately in Table 2.3.

PCR reaction components

Component	Final concentration
10X PCR Buffer	1X
dNTPs	200 µM of each dNTP
Forward Primer	0.5 µM
Reverse Primer	0.5 µM
HotStar Taq DNA polymerase	2.5 units/100 µl reaction
$MgCl_2$	1.5 mM
DNA template	
Distilled Water	

dNTPs deoxynucleotide triphosphates (Promega)

Amplification Parameters

	Step	Time	Temperature °C
1.	Initial Activation	15 min	95
Three step cycling			
2.	Denaturation	45 sec	94
3.	Primer Annealing	30 sec	~5°C below Tm of primers
4.	Extension	1 min	72
Number of cycles: 25-30			
5.	Final extension	10 min	72

2.2.8. Quantitative PCR

100ng gDNA was used per reaction. Taqman quantitative PCR for integrated GFP copies was performed on duplicate samples. For each sample, Taqman quantification of 18S rRNA copies was performed in parallel to control for input. The QuantiTect Probe PCR Master Mix (Qiagen) was used in Q-PCR reactions; the reaction components are shown in the table below. The sequences of GFP and 18S primers and sequence-specific probes are shown in Table 2.3.2. Quantitative PCR reactions were performed in a 25 µl volume in 96 well plates using the ABI PRISM 7000 or Eppendorf Masterplex real time thermal cyclers. In the reactions in which GFP copies were quantified, 0.02 µg/ µl Salmon Testis DNA (Sigma) was present. Absolute quantification of both GFP and 18S copies was performed using plasmid DNA standards. For GFP quantification, lentiviral or retroviral transfer vector plasmid was used. For 18S quantification, a 600bp fragment of the 18S rRNA gene was amplified using the primers 18S rRNA-F and 18S rRNA-R. A 10 fold serial dilution series of plasmid DNA, starting at 10^5 copies and going down to 10^1 copies, was amplified and used to generate a standard curve. The number of cells in each sample was calculated by dividing the amount of DNA in the sample in pg by 13.2 (13.2 pg is the DNA content of a tetraploid BAF3 cell). The cell number was normalised to the number of 18S copies measured in the sample, assuming the number of 18S copies per cell is a constant. To calculate the average number of GFP copies per cell, the number of GFP copies measured by Q-PCR was divided by the normalised cell number. For each sample the average number of GFP copies per cell was further normalised to that obtained for mutant HV A2. This mutant contains a single copy of the vector per genome and was included as a control in each Q-PCR run.

Q-PCR reaction components

	GFP	18S
Genomic DNA	2 µl (~100 ng)	2 µl (~100 ng)
Salmon Testis DNA	0.5 µg	-
Sequence-specific probe	0.3 µM	0.4 µM
Forward primer	0.3 µM	0.4 µM
Reverse primer	0.15 µM	0.2 µM
QuantiTect Probe PCR Master Mix	12.5 µl	12.5 µl
ddH2O	Up to 25 µl total volume	Up to 25 µl total volume

Amplification parameters

	Step	Time	Temperature °C
1.	Initial Activation	15 min	95
Two step cycling			
2.	Denaturation	15 sec	94
3.	Annealing/ Extension Data Acquisition	60 sec	60
Number of cycles: 40			

2.2.9. Integration site PCR

2.2.9.1. Inverse PCR

Wild-type HIV-1 and MLV vector integration sites were cloned out by inverse PCR (IPCR) using a protocol modified from that described by (Carteau *et al.* 1998). 2µg gDNA was digested overnight with NlaIII (HV only) or MspI (HV/CNCG/MFG). To remove enzymes and salts, reactions were cleaned up using the QIAquick PCR Purification kit (Qiagen). Digested DNA was eluted in EB buffer in a final volume of 50 µl and DNA concentration was measured by NanoDrop spectrophotometer. To obtain intramolecular ligation of individual fragments, 200 ng digested genomic DNA was ligated in 500 µl volume using T4 DNA ligase (Promega). The ligation was performed in a Peltier thermal cycler (Dyad) for 600 cycles of 30 sec at 12°C and 30 sec at 19°C. The ligation samples were cleaned using Microcon YM-30 columns (Amicon) and eluted in a final volume of 10 µl. 5 µl of this sample was amplified using 1st round PCR primers, shown in table 2.3.3. In nested amplification, 1 µl of first round product was used.

2.2.9.2. Linker mediated PCR

SIN lentiviral vector integration sites were cloned using LM-PCR, as described by (Wu *et al.* 2003). 1µg gDNA was digested overnight with NlaIII, TaqI or MspI restriction enzymes. A second digest, to destroy the internal band generated from the 5'LTR, was carried out for 2 hours with either EcoRI (for NlaIII or MspI initial digest) or SacI (for TaqI initial digest). Digested DNA was column purified using a PCR Purification Kit (Qiagen). To generate the linker cassette, a positive linker oligonucleotide was annealed to a negative linker oligonucleotide. 100pmol of each linker oligonucleotide was heated to 95°C for 5 minutes and then slowly cooled to room temperature. 0.5pmol of linker cassette was ligated overnight at room temperature to digested DNA in a total volume of 35µl. This reaction was again column purified using a PCR Purification Kit. One third of the eluate was used in the first round of PCR. PCR reactions were performed using the HotStarTaq DNA Polymerase (Qiagen). In nested PCR, 1µl of a 1:2 dilution of

first round product was used per 50µl PCR reaction. Sequences of linker oligonucleotides and primers used in LM-PCR are shown in the table 2.3.4.

PCR products were separated by gel electrophoresis on 2% agarose gels. Individual bands were cut out of the gel and purified using a gel extraction kit (Qiagen). Each PCR fragment was eluted in a final volume of 30 µl. 3 µl of this eluate was cloned into the pGEM T Easy Vector (Promega). Commercial sequencing of plasmids containing host-virus junctions was carried out by Lark Technologies.

2.2.10. Integration site mapping

Sequence from a genuine HV vector integration site obtained by inverse PCR had to contain: at one end, 26bp (NlaIII) or 41bp (MspI) of HIV-1 U3 region sequence from the HUC forward primer to the restriction site; and at the other end, 34bp of U3 sequence from the IP3 reverse primer to the beginning of the 5'LTR. Sequence from a genuine retroviral vector integration site obtained by inverse PCR had to contain: at one end, 73bp of MLV U3 region sequence from the Nested 5F forward primer to the MspI restriction site; and at the other end, 30bp of U3 sequence from the Nested 5R reverse primer to the beginning of the 5'LTR.

Genuine LM-PCR product sequences had to contain 3'LTR sequence from the nested primer to the end of the 3'LTR and also the linker sequence.

Vector flanking DNA sequence was blasted against the mouse genome, released August 2007 (www.ensembl.org/Mus_musculus). The same criteria as used by Wu *et al* were used to map vector-genome junction sequences onto the mouse genome (Wu *et al.* 2003). The sequence had to match a genomic location starting within 3bp after the beginning of the 5'LTR (inverse PCR) or the end of the 3'LTR (LM-PCR). The sequence had to show >95% identity to the genomic sequence and could only match one genomic locus with >95% sequence identity. Important integration sites were confirmed by locus-specific PCR. To amplify the 5'LTR junction, a locus specific forward primer was designed and used in combination with a U3 region specific reverse primer. To amplify the 3'LTR

junction, a locus specific reverse primer was designed and used in combination with a U5 region specific forward primer. Primer sequences are shown in Table 2.3.5. Some of these locus-specific PCR products were sequenced. It was checked whether 5bp (HIV) or 4bp (MLV) duplications of host DNA flanked the 5' and 3' LTRs.

2.2.11. Multiplex PCR

In experiment LV3, multiplex PCR was used to identify additional HV vector integration sites in the Ghr locus. Seven different forward primers were designed every 2kb in the 15kb region immediately upstream of Ghr exon 2. A 10x primer mix was prepared containing 2µM of each locus-specific forward primer and 4µM of the HV vector-specific IP3 reverse primer. Multiplex PCR was performed on 150ng gDNA using a Multiplex PCR Kit (Qiagen). Cycling conditions were as follows: 95°C 15'; 30x (94°C 30", 55°C 90", 72°C 2'); 72°C 10'. Sequences of locus-specific forward primers are shown in Table 2.3.6.

2.2.12. RNA extraction

Cells were grown in T75 flasks to a density of 5 x 10^5 cells/ml. 2 x 10^7 cells were pelleted in a 50 ml Falcon and washed twice in HBSS. Cells were lysed by repeatedly pipetting them in 1ml of RNAzol B reagent (Biogenics). The lysate was either frozen at -80°C for RNA extraction at a later date; or extraction was proceeded with immediately. To separate the aqueous from the organic phase of the lysate, 200µl Chloroform was added per sample, vortexed and kept on ice for 15min. The homogenate was then centrifuged at 4°C for 15 minutes at 13.000 rpm. Next, RNA was precipitated from the upper aqueous phase by adding an equal volume of isopropanol. This was left on ice for 15 minutes and then centrifuged at 4°C for 30 minutes at 13.000 rpm. The RNA pellet was washed twice with 200µl 70% EtOH (in DEPC-treated water). The RNA pellet was dissolved in 40µl RNase-free water, RNA concentration was measured and RNA was stored in aliquots at -80°C.

RNA for use in Q-RT-PCR was extracted from 5 x 10^6 cells using an RNeasy kit (Qiagen). Cells were washed twice in HBSS. After the final wash, cells were re-

suspended in a denaturing guanidine-thiocyanate–containing buffer (buffer RLT) containing β-mercaptoethanol. The lysate was frozen at -80°C for at least one hour. The lysate was thawed on ice and ethanol was then added. This sample was then applied to an RNeasy column and washed twice. RNA was eluted in 50 µl RNase-free water.

2.2.13. Reverse-transcriptase PCR

The reverse transcription reaction was performed using a Protoscript First Strand cDNA synthesis kit (New England Biolabs). First strand cDNA from Ghr, Igf1 or Il3 mRNAs was synthesised using 2 µg total RNA as a template. The RT reaction was primed using random primers (Promega) or a gene specific reverse primer (Shown in Table 2.3.7). Firstly, RNA, 20pmol gene-specific primer/30pmol random primer and dNTPs in a total volume of 16 µl were heated for 5 minutes at 70°C; next 2 µl 10X RT Buffer, 1 µl Mo-MLV reverse transcriptase (RT) and 1 µl RNase inhibitor were added per reaction. Reactions without RT were also performed on every sample. The samples were then incubated for 1 hour at 42°C. The enzyme was inactivated for 5 minutes at 95°C. The reaction was diluted to a total volume of 50 µl. 2 µl of this reaction was used as template in the subsequent PCR reaction. Primer sequences are shown in Table 2.3.7. RNA from mouse liver was used as a positive control in the Ghr and Igf1 RT PCR reactions; WEHI3B RNA was used as a positive control in Il3 RT PCR.

2.2.14. Quantitative RT-PCR

cDNA was generated from 1µg RNA using the QuantiTect Reverse Transcription kit (Qiagen). This kit contains a genomic DNA wipe-out reaction and uses random primers and oligo d(T) to prime the RT reaction. One eighth of this reaction (2.5µl) was used in Taqman quantitative PCR using the QuantiTect Probe PCR Master Mix (Qiagen). Reaction components are shown in the table below; amplification parameters were as described previously. The IL-3 primers and probe were designed using PrimerExpress software. The sequences are shown in Table 2.3.2. The dual labelled IL-3 probe spans the junction of exons 4 and 5. pGEM T Easy vector containing the complete IL-3 cDNA as an insert was

used for absolute quantification of IL-3 copies. A 10 fold serial dilution series of this plasmid, starting at 10^5 copies and going down to 10^1 copies, was amplified and used to generate an IL-3 standard curve. Taqman quantification of 18S rRNA copies was performed in parallel, as described previously. A 10 fold serial dilution series of the 18S plasmid, starting at 10^{10} copies and going down to 10^6 copies, was amplified and used to generate an 18S standard curve. The ratio of IL-3 copies per 18S copies was calculated from duplicate samples.

Q-RT-PCR reaction components

	IL-3	**18S**
RT reaction	2.5 µl	2.5 µl
Sequence-specific probe	0.3 µM	0.4 µM
Forward primer	0.3 µM	0.4 µM
Reverse primer	0.15 µM	0.2 µM
QuantiTect Probe PCR Master Mix	12.5 µl	12.5 µl
ddH2O	Up to 25 µl	Up to 25 µl

2.3. Appendix

Table 2.1. Buffers and solutions

1X PBS (Phosphate-buffered saline)	1.9mM NaH2PO4; 8.1mM Na2HPO4; 154mM NaCl; pH 7.4
1X HBSS (Hanks balanced salt solution)	5.4mM KCl; 0.3mM $Na_2HPO_4 \cdot 7H_2O$; 0.4mM KH_2PO_4; 4.2mM $NaHCO_3$; 1.3mM $CaCl_2$; 0.5mM $MgCl_2 \cdot 6H_2O$; 0.6mM $MgSO_4 \cdot 7H_2O$; 137mM NaCl; 5.6mM D-glucose; pH 7.4
1X TAE buffer	40mM Tris (pH 7.8), 20mM sodium acetate and 1mM EDTA

Elution buffers	
TE	10 mM Tris.Cl, 1mM EDTA, pH 8.0
EB	10 mM Tris-Cl, pH 8.5
AE	10 mM Tris-Cl, 0.5 mM EDTA, pH 9.0

Bacterial media and solutions	
LB (Luria-Bertani) agar	1% bacto typtone, 0.5% bacto yeast, 0.5% NaCl (pH 7.0) with 15g/L bactoagar
LB (Luria-Bertani) broth	1% bacto typtone, 0.5% bacto yeast and 0.5% NaCl (pH 7.0)
SOC medium	2% tryptone; 0.5% yeast extract; 10 mM NaCl; 2.5 mM KCl; 10 mM $MgCl_2$; 10 mM $MgSO_4$; 20 mM glucose
TFB-I	30 mM potassium acetate; 100 mM rubidium chloride. 10 mM calcium chloride; 50 mM magnesium chloride. 15% glycerol; pH 5.5 with acetic acid
TFB-II	10 mM MOPS; 75 mM calcium chloride; 10 mM rubidium chloride; 15% glycerol; pH 6.5 with KOH

Genomic DNA extraction buffers	
Lysis buffer	10mM Tris.Cl (pH 7.4); 10mM EDTA; 10mM NaCl; 0.5 % SLS; 1mg/ml Proteinase K
Precipitation buffer	18.75µl of 5M NaCl/1250µl Ethanol 75mM

Southern Blotting solutions	
Depurination Solution	0.25M HCl
Denaturation Solution	0.5 M NaOH 1.5 M NaCl
Neutralisation Solution	1 M Tris Cl 1.5 M NaCl pH 7.4
20X SSC	3.0M NaCl 0.3M sodium citrate pH 7.0
Wash Solution 1	2X SSC 0.1X SDS
Wash Solution 2	0.1X SSC 0.1X SDS

Table 2.3. Primer Sequences

Table 2.3.1. Primers used to generate Southern Blot probe

GFP-F	ATG GTG AGC AAG GGC GAG GA
GFP-R	TTA CTT GTA CAG CTC GTC CAT GCC

Table 2.3.2. Quantitative (RT) PCR Primer and Probe Sequences

GFP	
GFP-TQ-F	CAA CAG CCA CAA CGT CTA TAT CAT
GFP-TQ-R	ATG TTG TGG CGG ATC TTG AAG
GFP-TQ-P	5'-FAM-CCG ACA AGC AGA AGA ACG GCA TCA A-3'TAMRA
18S	
18S-TQ-F	TCG AGG CCC TGT AAT TGG AA
18S-TQ-R	CCC TCC AAT GGA TCC TCG TT
18S-TQ-P	5'-FAM-AGT CCA CTT TAA ATC CTT-3'TAMRA
IL-3	
IL-3-TQ-F	AAC TTA ACT GTT GCC TGC CTA CAT C
IL-3-TQ-R	AAG TCA TCC AGA TCT CGA ATG AAG AC
IL-3-TQ-P	5'-FAM-C GAA TGA CTC TGC GCT GCC AGG G-3'TAMRA
Primers used to generate 18S and IL-3 inserts for plasmid DNA standards	
18S-F	TAC CTG GTT GAT CCT GCC AGT A
18S-R	TTA CGA CTT TTA CTT CCT CTA GAT AG
IL-3-F	ATG GTT CTT GCC AGC TCT ACC
IL-3-R	TAG ATA AAT TGA TAA GAC ATT TGA TGG C

F = forward primer; R = reverse primer; P = probe
All primers were synthesised by Invitrogen; 5'-FAM/3'-TAMRA dual labelled probes were synthesised by Sigma.

Table 2.3.3. Inverse PCR primer sequences and cycling conditions

Lentiviral inverse PCR (1st round) 95°C 15'; 35x (94°C 45", 53°C 30", 72°C 1'); 72°C 10'	
HUE	AGC CAA TGA AGG AGA GAA CAC CC
HUB	GAT CAA GGA TAT CTT GTC TTC GT
Lentiviral inverse PCR (Nested) 95°C 15'; 30x (94°C 45", 54°C 30", 72°C 1'); 72°C 10'	
HUC	CTT GTT ACA CCC TGT GAG CCT
IP3	TCT TGT CTT CGT TGG GAG TGA
Retroviral inverse PCR (1st round) 95°C 15'; 35x (94°C 45", 51°C 30", 72°C 1'); 72°C 10'	
MspI-5F	GAG AAG TTC AGA TCA AGG TC
MspI-5R	TCC ATG CCT TGC AAA ATG GC
Retroviral inverse PCR (Nested) 95°C 15'; 30x (94°C 45", 53°C 30", 72°C 1'); 72°C 10'	
Nested-5F	CAA GGT CAG GAA CAG ATG GA
Nested-5R	GCT TGC CAA ACC TAC AGG TG

Table 2.3.4. Linker mediated PCR primer sequences and cycling conditions

LM PCR Linker Plus	
NlaIII Plus	GTA ATA CGA CTC ACT ATA GGG CTC CGC TTA AGG GAC TAC ATG
TaqI/MspI Plus	GTA ATA CGA CTC ACT ATA GGG CTC CGC TTA AGG GAC
LM PCR Linker Minus	
NlaIII Minus	PO_4-TA GTC CCT TAA GCG GAG
TaqI/MspI Minus	PO_4-GC GTC CCT TAA GCG GAG
CSGW LM PCR (1st round) 95°C 5'; 30x (94°C 45", 55°C 30", 72°C 1'); 72°C 10'	
Linker 1	GTA ATA CGA CTC ACT ATA GGG C
HIV-1 3'LTR 1	AGT GCT TCA AGT AGT GTG TGC C
CSGW LM PCR (Nested) 95°C 5'; 30x (94°C 45", 55°C 30", 72°C 1'); 72°C 10'	
Linker Nested	AGG GCT CCG CTT AAG GGA C
HIV-1 3'LTR Nested	GTC TGT TGT GTG ACT CTG GTA AC

Table 2.3.5.a.Sequences of site-specific PCR primers for confirmation of lentiviral vector integration sites

Experiment Mutant Name	**Confirmation of 5'LTR junction:** Locus-specific forward primer (Used in combination with HIV-1 U3 primer **IP3** TCT TGT CTT CGT TGG GAG TGA)	**Confirmation of 3'LTR junction** Locus-specific reverse primer (Used in combination with HIV-1 U5 primer **GT250** GAC CCT TTT AGT CAG TGT GGA A)
LV1 HV A2	HV A2-F: GAG CTA CAA GGT CAT ACA CTC	HV A2-R: GAG CAC AGA GGA AGG TGG AAA
LV2 HV3	HV3-F: GGA TTC AGT ATC TGG AAC TTC	HV3-R: AAT GCA TCA ATA CAT AGA CTG
LV2 HV14	HV14-F: GGG ATC AGA GCT AAA CAA AGA	HV14-R: CCA AGT CCT AGC CAA CAT TCT
LV3 HV32/HV49	HV32/HV49-F: CAG TAA CAT CCC ACT TCC TGA	HV32/HV49-R: GCT CAG CAC AGG GCT TTG AAT
LV3 HV41	HV41-F: CGC TGC TTC AGT CCT GTT CTA CCT	HV41-R: TTT CCA CAC TCC AAG TCT GAG GCA
LV3 HV43	HV43-F: GTT CAG TGT TTG AGA GGA CCT CTG	HV43-R: GTG TCG GAG TTG ACT GAA GTT ACC
LV4 GH CSGW1-1	Zfp407-F: GAG CAGA AGA CCA ACA AAG CA	ND
LV4 GH CSGW1-1 and GH CSGW2-4	Cnot4-F: GAA TGG TGA TAT CCC TGC CTT	ND
	Gmeb-2-F: GCA AGT GAT TAC CTC TGC TCA	ND
	Wipi2-F: TTA GGT GTG GGC ACA GGT TGT	Wipi2-R: CTA AAC TGA AAC CTC AGC CCT

Table 2.3.5.b. Sequences of site-specific PCR primers for confirmation of retroviral vector integration sites

Mutant	**Confirmation of 5'LTR junction:** Locus-specific forward primer (Used in combination with MLV U3 primer **MspI5R** TCC ATG CCT TGC AAA ATG GC)	**Confirmation of 3'LTR junction** Locus-specific reverse primer (Used in combination with MLV U5 primer **MspI3F** CCT CTG AGT GAT TGA CTA CC)
C40	C40-F: CAC TAG TCA GGT GGG ATG AAG T	C40-R: TAA GCA GAC CCT TAC CTC CGT C
C57 AA (Chr 11)	C57 AA-F: ACT CTT GGA CTT CCT ACT CAG G	C57 AA-R: TGT AGG CAG CCA GTG TCA AGA G
C57 AB (Chr 7)	C57 AB-F: CGC ATC ACT TCC TGC TAG GTT G	C57 AB-R2: CTC AGG ACT CTC CAT ATC TGC T
C57 CA (Chr 9)	C57 CA-F: CAG GGA CTA CAC AGT GAG ATC C	C57 CA-R: AGG GCC TCA CAA ATG TGT TCC T
C57 DA (Chr X)	C57 DA-F2: CAG TCT GCA GAC ACA AGT GAA G	C57 DA-R2: ACC TGT GGA ATC TGA TGG TGA G
C96 BC (Chr 7)	C96 BC-F: ATA GGG AGT TGA GGT GGG GGA A C96 BC-F2: TGG CTG GGT GAC ATG ATG AGC T	C96-BC-R: CCC TTT GTG TGG TTC CCA CCA T C96-BC-R2: GAA AGA CCA GCT CTC ACT CAC A
G18 AA (Chr 11)	G18 AA-F: ATC TGG AGA CAG TGC TAA CCT C	G18 AA-R: AAA GCC CTG TCA TGG GTC ATC T
G18 AD (Chr 9)	G18 AD-F: TGTCAGTGGCCCTGGTAG GATA	G18 AD-R: GCTTGCTCAGCCAACTTTC TCA

Table 2.3.6. Sequences of Multiplex PCR primers

Name	Locus-specific forward primer (Used in combination with HIV-1 U3 primer **IP3**)
108-HV-F	TGC TCA GGT TGT CTC AGA GTA
110-HV-F	GGA GGC AAA GCC TCT TGA TAA
112-HV-F	GCT AAG CTT GGA TTT ATG CGC T
114-HV-F	GGC CAA CGT TTG ATG GAG TAA T
116-HV-F	GGT AGA CTT TAG AGG TAC TGA G
118-HV-F	GTT CAT GCT AAG GTG AGT GAA G
120-HV-F	AGC CTT AGT ATA CAG TGG GCT A

Table 2.3.7. Reverse – transcriptase PCR primers

Name	Application	Sequence	Product size
Growth Hormone Receptor			
GHR A1	F-primer Exon 4	TGG ACA GAA GGA GAT AAT CCT GAT	552
GHR RT3	R-primer Exon 8	GAA CTC GCT GTA CTT TTC AAA GCT C	
Interleukin 3			
IL3 FW	F-primer Exon 1	ATG GTT CTT GCC AGC TCT ACC	602
IL3 RW	R-primer Exon 5	TAG ATA AAT TGA TAA GAC ATT TGA TGG C	
Insulin-like growth factor 1			
Igf1-F	F-primer Exon 2	TCG TCT TCA CAC CTC TTC TAC CTG	321
Igf1-R	R-primer Exon 3	CTT CTG AGT CTT GGG CAT GT	

Chapter 3

3. Optimisation of mutagenesis assay conditions

3.1. Introduction

In the mutagenesis assay I tried to establish, IL-3 independent mutants of an IL-3 dependent cell line, BAF3, were selected. The BAF3 cell line was established in the early 1980s from mouse bone marrow of BALB/c mice. It is pro-B cell in phenotype and its growth is dependent on IL-3 (Palacios *et al.* 1984). The BAF3 cell line was chosen instead of other cytokine-dependent cell lines for the purpose of setting up an insertional mutagenesis assay for several reasons. Firstly, BAF3 cell biology is well characterised. The signalling pathways activated in this cell line downstream of the IL-3 receptor are known. Other cytokines and their receptors, as well as oncogenic proteins, that can cause IL-3 independent survival and proliferation of BAF3 cells have been described. This knowledge should be helpful in identifying vector induced transforming events in any BAF3 cell mutants.

Secondly, BAF3 cells were reported to undergo apoptosis reliably upon IL-3 withdrawal. This is a pre-requisite for a cell line in which factor independent mutants are selected following vector transduction. Lastly, retroviral vectors had previously been used as insertional mutagens of the BAF3 cell line in our lab with the purpose of identifying anti-apoptotic genes. These studies, though not quantitative, offered a proof of principle that retroviral vectors could transform these cells and a single transforming event was sufficient to do so.

The above points will be discussed in greater detail in the sections that follow. I will start with a review of the signalling pathways activated by IL-3 in BAF3 cells. Next, the mechanisms by which BAF3 can be transformed to IL-3 independence will be discussed. Loosely based on the conditions of the earlier retroviral insertional mutagenesis studies in BAF3 cells, one pilot lentiviral vector mutagenesis experiment had already been performed before I joined the lab. The results section of this chapter will describe optimisation of these assay conditions.

3.1.1.IL-3 signalling

The following reviews on IL-3 signalling were consulted during the writing of this section: (Reddy *et al.* 2000; Baker *et al.* 2007; Robb 2007)

3.1.1.1. IL-3 and the IL-3 receptor

The haematopoietic cytokines IL-3, IL-5 and GM-CSF are encoded by mouse chromosome 11. IL-3 promotes survival and proliferation of haematopoietic progenitor cells. In vivo IL-3 is produced under conditions of stress, mainly by activated T cells to stimulate haematopoiesis and recruit more differentiated immune cells to the periphery. The targets of IL-3 are progenitor cells of both the lymphoid and myeloid lineages; the IL-3 receptor is expressed on the cell surface of these cells (Robb 2007). The IL-3 receptor belongs to the gp140 family of class I cytokine receptors (Reddy *et al.* 2000). It is a heterodimeric receptor consisting of a unique IL-3 specific α chain and a common β chain (β_c/ gp140/ CD131). β_c is shared by the IL-5 and GM-CSF receptors; it is responsible for intracellular signal transduction by these receptors. Mice also encode an IL-3 specific β chain (β_{IL-3}) in addition to β_c. IL-3 receptor dimerisation only occurs once IL-3 has bound to the α chain; the β_c is then recruited to the α chain and the two receptor subunits become linked through disulphide bonds, causing intracellular signalling to be initiated.

3.1.1.2. IL-3 receptor intracellular signalling pathways

IL-3 mediates the survival and proliferation of haematopoietic progenitor cells through the activation of three known signalling cascades that eventually result in activation of gene expression in these cells. The way in which aberrations in signalling by these pathways can result in transformation of IL-3 dependent BAF3 cells will be discussed in the next section.

The IL-3 receptor does not itself have any catalytic activity. Janus kinases (JAKs) are at the heart of all three signalling pathways activated by IL-3 (Baker *et al.* 2007). Upon IL-3 receptor dimerisation, JAK2 becomes activated by autophosphorylation. Activated JAK2 then goes on to phosphorylate the intracellular domain of β_c on 6 critical tyrosine residues (Y577, Y612, Y695,

Y750, Y806 and Y866 (Martinez-Moczygemba and Huston 2003)). These phosphotyrosines serve as docking sites for many different signalling molecules that bind via their SH2-domains, including Signal Transducers and Activators of Transcription (STATs), Src family kinases, protein phosphatases (needed for signal termination) and adaptor molecules such as Shc, Grb2 and Phosphatidylinositol 3 kinase (PI3K)(Baker *et al.* 2007). Src-family kinases fyn, hck and lyn also play an important role in IL-3 signal transduction. These three kinases where shown to become activated upon IL-3 receptor dimerisation, however it is not clear whether this happens at the same time as JAK2 phosphorylation or as a downstream event of JAK2 activation (Corey and Anderson 1999).

3.1.1.2.1. Signal Transducers and Activators of Transcription (STAT) signalling pathway

The STAT signalling pathway is activated by most cytokine receptors. STAT family members STAT3, 5a and 5b are activated by the IL-3 receptor. STATs are a group of transcription factors that are found in an inactive state in the cytoplasm. Three phosphotyrosines (Y612, Y695 and Y750) in the membrane proximal region of β_c serve as docking sites for STATs. β_c bound STATs become tyrosine phosphorylated on their carboxy termini, which is mediated by JAK2 or Src family kinases. Phosphorylated STATs dimerise and translocate to the nucleus where they function as active transcription factors. STAT5 target genes include Bcl-X_L, Bcl-2, c-myc, cyclin D1/D2/D3/E and pim-1 (Baker *et al.* 2007). STAT5 is a mediator of IL-3 induced cell survival and proliferation. BAF3 expressing dominant negative STAT5 are partially resistant to IL-3 induced cell proliferation. They are also more sensitive to apoptosis, because they express lower levels of Bcl-X_L (Mui *et al.* 1996; Dumon *et al.* 1999). STAT5 was also shown to mediate NF-kappaB activation by IL-3 in BAF3 cells (Nakamura *et al.* 2002). NF-kappaB was shown to be necessary for IL-3 mediated cell survival (Besancon *et al.* 1998).

3.1.1.2.2. Ras - Mitogen Activated Kinase Pathway

The adaptor molecule Shc is another molecule recruited to a phosphotyrosine (Y577) on β_c. It becomes phosphorylated and associates with another adaptor molecule Grb2. This complex of adaptor molecules subsequently associates with the guanine nucleotide exchange factor (GNEF) mSos which exchanges GDP for GTP on Ras. This leads to activation of the Raf-1/ MEK/ ERK pathway. This MAPK cascade eventually results in the activation of c-fos and c-jun; two transcription factors that heterodimerise with each other and bind to AP-1 sites that are present in the promoters of genes that promote cell proliferation.

ERK phosphorylates the pro-apoptotic protein BAD on serine 112 (Scheid *et al.* 1999). (Akt phosphorylates BAD on serine 136, see next section). Dually phosphorylated BAD is sequestered in the cytosol by 14-3-3 and hence is inactive. Unphosphorylated BAD promotes cell death by binding and inhibiting the activity of Bcl-X_L and Bcl-2 (del Peso *et al.* 1997), so that the pro-apoptotic protein Bax is displaced. Bax dimers induce apoptosis by associating with the mitochondrial membrane, leading to cytochrome c release and caspase activation (Low *et al.* 2001).

3.1.1.2.3. Phosphatidylinositol 3 kinase (PI3K) Pathway

Upon IL-3 signalling, serine 585 on β_c is phosphorylated by protein kinase A. 14-3-4 docks here and in turn binds PI3K. PI3K, by phosphatidylinositol triphosphate (PIP_3) generation, recruits and activates Akt serine/ threonine kinase. Akt itself inactivates a number of pro-apoptotic proteins by phosphorylation. These include BAD (on serine 136 (del Peso *et al.* 1997), see previous paragraph), caspase 9 and FKHRL1 (Brunet *et al.* 1999). This latter protein is a transcription factor involved in the transcription of pro-apoptotic genes such as Fas ligand. It localises in the cytoplasm after phosphorylation. PI3K also stimulates glycolysis; maintenance of ATP levels is essential to cell survival (Mathieu *et al.* 2001).

3.1.1.3. Why do BAF3 cells undergo apoptosis upon IL-3 withdrawal?

Upon IL-3 withdrawal there is expected to be much reduced IL-3 receptor signalling. In BAF3 cells downregulation of Bcl-X_L, but not Bcl-2, mRNA and protein occurs upon IL-3 withdrawal (Leverrier *et al.* 1997). Existing Bcl-X_L protein is sequestered by BAD, releasing its pro-apoptotic counterpart Bax (Low *et al.* 2001).

3.1.2. BAF3 transformation to IL-3 independence

3.1.2.1. Over-expression of anti-apoptotic proteins

Insertional activation by retroviral vectors of the Bcl-X gene in BAF3, resulted in a mutant that survives but does not proliferate in the absence of IL-3 (Thomas *et al.* 1998). Over-expression of another anti-apoptotic protein Bcl-2 in BAF3 gave a similar phenotype (Collins *et al.* 1992). In order for this type of mutant to proliferate in the absence of IL-3, c-myc expression is also required. It was shown that both Bcl-2 and Bcl-X_L in cooperation with c-myc could induce cytokine independent growth of BAF3 (Malde and Collins 1994; Miyazaki *et al.* 1995; Nosaka *et al.* 1999).

3.1.2.2. Aberrant expression of growth factors or their receptors

BAF3 cells are dependent on IL-3. One obvious way in which BAF3 can become IL-3 independent is by autocrine production of this cytokine. Cytokines and growth factors other than IL-3 could also provide survival and proliferation signals substituting those of IL-3 if these factors are produced in an autocrine fashion. Asides from IL-3, BAF3 were shown to be responsive to IL-4 and IGF-1, but only IGF-1 could sustain long term survival (Leverrier *et al.* 1997).

Alternatively, expression of growth factor receptors not normally expressed by BAF3 cells could confer IL-3 independence provided their ligands are present in the selection medium. BAF3 constitutively express the receptors for IL-3, IL-4 and IGF-1. The IL-3 and IL-4 receptors are made up of several subunits. Complementation of the β_c (of **IL-3**/IL-5/GM-CSF) and γ_c (of IL-2/**4**/7/9/15/21) with other α chains could make BAF3 responsive to other cytokines. This has

been demonstrated when the complete cDNAs of the receptors for IL-2 (Collins *et al.* 1990), IL-5 (Pless *et al.* 1997), GM-CSF (Rosas *et al.* 2007), IL-9 (Demoulin *et al.* 1996), IL-15 (Bulanova *et al.* 2003) and IL-21 (Ozaki *et al.* 2000) were introduced into BAF3. We are planning to select in serum only without additional cytokines, so this type of mutant would not be picked up.

Receptors for cytokines of the growth hormone family such as erythropoietin (EPO), thrombopoietin (TPO), G-CSF, prolactin (PRL) and growth hormone (GH) consist of a single subunit that homodimerises upon ligand binding. BAF3 transfected with the EPO-R (Jones *et al.* 1990), MPL (Alexander *et al.* 1995) , G-CSF-R (Li and Sartorelli 1995), PRL-R (Sasaki *et al.* 1996) and GHR (Jeay *et al.* 2000) are still cytokine-dependent, but whereas they could previously only survive and proliferate in response to IL-3, they can now also do so in response to EPO, TPO, G-CSF, PRL and GH, respectively. Only constitutively active mutants of these receptors confer complete cytokine-independence to BAF3. For example, an MPL mutant with a single amino acid substitution in the transmembrane domain that results in constitutive receptor dimerisation has been described. BAF3 cells expressing this mutant MPL are independent of any cytokine and even serum for their growth (Alexander *et al.* 1995; Onishi *et al.* 1996). Epidermal growth factor (EGF), fibroblast growth factor (FGF) and platelet derived growth factor (PDGF) and growth hormone (GH) are present in serum. EGF can support short term proliferation of EGFR expressing BAF3 cells (Collins *et al.* 1988). Expression of FGFR1 can confer FGF-dependent growth onto BAF3 cells (Wang *et al.* 1994) and expression of PDGFRβ can confer PDGF-dependent growth onto BAF3 cells (Irusta and DiMaio 1998).

3.1.2.3. Aberrant expression or activity of intracellular signalling proteins

BAF3 cells are potently transformed by constitutively active protein kinases as described below. However, over-expression of these proteins per se for example by insertional activation, may not be sufficient to transform these cells.

The three main signal transduction pathways activated by IL-3 are the STAT, Ras-MAPK and PI3K/Akt pathways.

3.1.2.3.1. JAK2

The JH2 domain of JAK negatively regulates the activity of this protein. JH2-deficient JAK2 is itself hyperphosphorylated and leads to increased STAT5 phosphorylation. JH2-deficient JAK2 leads to IL-3 independent proliferation of BAF3 cells (Malinge *et al.* 2007).

3.1.2.3.2. STAT

It was shown that the introduction of a constitutively active STAT5 mutant into BAF3 cells conferred cytokine-independent proliferation to these cells and also protected against DNA-damage induced apoptosis (Onishi *et al.* 1998; Hoover *et al.* 2001). Onishi et al used a STAT5 double point mutant that was hypertyrosine phosphorylated. BAF3 expressing this mutant proliferated independently of IL-3 but were still dependent on serum. Introduction of this STAT5 mutant in combination with a constitutively active Raf-1 mutant resulted in an IL-3 and serum-independent mutant. Of note, introduction of wild type STAT5 into BAF3 never gave rise to IL-3 independent growth. Constitutively active STAT5 was shown to increase expression of Bcl-X_L, pim-1 and c-myc.

3.1.2.3.3. Ras-MAPK

BAF3 cells expressing constitutively active Ras were able to proliferate independently of IL-3 (Hoover *et al.* 2001).

3.1.2.3.4. PI3K/Akt

BAF3 cells expressing myristoylated Akt showed increased viability but did not proliferate in the absence of IL-3. Mathieu et al describe a BAF3 mutant obtained by retroviral insertional mutagenesis in which Akt was found to be constitutively active. However, the retroviral vector integration site in this mutant was not specified. This mutant survived as well as proliferated in the absence of IL-3 (Mathieu *et al.* 2001).

3.1.2.4. Transformation by viral oncoproteins

Viral oncogenes (v-onc) are transduced by acutely transforming retroviruses; most subtly differ from their cellular proto-oncogene counterparts (c-onc) in that they carry mutations that confer constitutive enzymatic activity to viral oncoproteins (Also see section 1.2.1. of the introduction). *v-abl*, the oncogene carried by the Abelson murine leukaemia virus (A-MLV), was one of the first genes shown to transform BAF3 cells to IL-3 independence (Mathey-Prevot *et al.* 1986). *v-src* and *v-fps*, the v-onc of the Rous and Fujinami sarcoma viruses respectively, were also tested but they caused partial and no IL-3 independence of BAF3, though they did do so in other IL-3 dependent cell lines (reviewed in (Blalock *et al.* 1999)). Most of the v-onc found to cause factor-independence in IL-3 dependent cell lines are tyrosine kinase proteins that act upstream of the Ras/MAPK pathway.

3.1.2.5. Which type of mutant are we most likely to select for?

Essentially four types of potential BAF3 mutants have been described in this section. The first type of mutant has enhanced cell viability due to over-expression of anti-apoptotic proteins, but will not proliferate in the absence of IL-3. Whether this type of mutant is detected depends on the duration of IL-3 starvation. Cell viability is maintained for a few days, but eventually cells die anyway. IL-3 would have to be added back to the culture before this decrease in cell viability sets in, to allow this type of mutant to grow out. A second type of mutant is self-sufficient in cytokine production. These autocrine mutants should be able to grow out. Bystander cells may be rescued along the way via a paracrine effect if the cytokine is secreted to high enough levels. In a third type of mutant, IL-3 dependence has been substituted by dependence on another cytokine or growth factor whose receptor is not normally expressed on the cell surface of BAF3 cells. Detection of these mutants would dependent on the presence of the cytokines at sufficient concentrations in the selection medium. Lastly, mutants that are completely independent of any cytokines or even serum could be obtained. So far these mutants were seen only when constitutively

active protein kinases were introduced into BAF3 cells. Simple over-expression of key protein kinases alone would probably not be sufficient to transform BAF3 cells and additional mutations in these proteins are likely to be required. However, truncated proteins missing negative regulatory domains caused by vector insertion could potentially transform BAF3 cells.

3.1.3. Retroviral insertional mutagenesis in factor-dependent cell lines

Wild type retroviruses and retroviral vectors have been used to generate factor-independent mutants in cytokine dependent haematopoietic cell lines other than BAF3 (Table 3.1). The objective of these studies was to identify genes involved in the acquisition of growth factor independence and transformation of haematopoietic cells in leukaemia.

Factor independent mutants of the promyelocytic D35 cell line, that is normally dependent on IL-3 or GM-CSF for its growth, were described (Stocking *et al.* 1988). Long term marrow cultures of C3H/HeJ mice were immortalised with wild type Friend SFFV and Rauscher MLV retrovirus to establish the D35 cell line (Greenberger *et al.* 1980). The D35 cell line contains multiple (~10) copies of these viruses, which are also released from the cell. Factor-independent D35 mutants were obtained at a cell frequency of 2.4×10^{-7}. The mechanisms whereby these mutants had become growth factor independent were analysed (Stocking *et al.* 1988). 9 out of 11 mutants were found to secrete a growth factor that could support growth of parental D35 cells to a greater or lesser extent.

GM-CSF was produced by 6 mutants, and IL-3 by another 3 mutants. Provirus insertions were found 250-750bp upstream of the GM-CSF cap site in 3 out of 6 GM-CSF producing mutants. All three insertions were in reverse orientation to GM-CSF gene transcription. An intracisternal A particle (IAP) insertion in the 3'UTR was found in a fourth GM-CSF producing clone. In a follow-up study, an IAP insertion upstream of IL-3 was found in 1-2 IL-3 producing mutants (Heberlein *et al.* 1990). The two D35 mutants that did not produce a growth factor were found to express a common β chain of which the extracellular domain was truncated (Hannemann *et al.* 1995). No provirus/IAP insertions in this locus were

found. In both mutants, a large portion of mouse chromosome 15 was found to be deleted which encoded for both β_{IL-3} and the extracellular portion of β_C. These mutants are most likely spontaneous mutants that have arisen through chromosomal rearrangement. In addition, introduction of truncated β_C alone was not sufficient to transform D35 parental cells and secondary mutations were shown to be required (Prassolov *et al.* 2001).

Factor independent mutants of two IL-3 dependent cell lines, FDC Pmix and FDC-P1(M), were analysed. Both these cell lines had been immortalised by retroviral infection, but contained fewer proviral copies than D35. This explains the lower cell frequency at which IL-3 independent mutants were obtained. A number of IL-3 producing mutants were shown to have IAP insertions upstream of IL-3.

Insertional mutagenesis using retroviral vectors was performed in the human erythroleukaemia cell line TF-1, which is dependent on GM-CSF (Stocking *et al.* 1993). This study is interesting because it provides numbers at which factor-independent mutants were obtained. In total 1.58×10^8 TF-1 cells were transduced with a MPSV retroviral vector. Cells contained an average of 6.5 provirus copies, determined by quantitative Southern blot. 241 mutants were obtained at an integrant frequency of 2.3×10^{-7}. Unlike the study in D35 cells, only a minority of these mutants (13%) secreted a growth factor. No vector integration sites were analysed in this study. In a follow-up study, the phosphorylation status of signalling proteins in 14 mutants that did not produce growth factors was analysed. 11 out of 14 mutants were found to express constitutively phosphorylated $p60^{c\text{-}Src}$, but a possible relationship to retroviral vector integration site was never established (Horn *et al.* 2003).

The same MPSV retroviral vector was used in another study to obtain GM-CSF dependent mutants of the IL-3 dependent cell line FDC-P1(M)(Laker *et al.* 2000). This cell line only expresses the IL-3 receptor. Following vector transduction, cells were selected in serum plus GM-CSF to select for mutants that had insertionally activated the GM-CSF α chain. Over 100 mutants were obtained at an integrant frequency of 1.4×10^{-7}. 18 mutants were analysed and all expressed the GM-CSF α chain. In 40% gross rearrangements of the GM-CSF α chain locus were found, but neither provirus, LTR nor IAP insertions were present in these rearranged loci. Thus the role of the vector in generating these mutants could not be demonstrated.

In a follow-up study using the same vector and cell line, vector transduced FDC-P1(M) cells were subjected to an unbiased selection in medium containing serum only, without additional cytokines (Meyer *et al.* 2002). Mutants were obtained at an integrant frequency of 1.5×10^{-7}. ~10% of mutants secreted a cytokine. In one mutant, a vector integration site 1kb upstream of Pdgfrbeta was found. This mutant expressed the PDGF receptor. Although, no activating mutations in the receptor itself could be detected, over-expression of this gene in parental FDC-P1(M) was not sufficient to confer IL-3 independence to these cells. It was therefore postulated that secondary mutations had contributed to IL-3 independence in the Pdgfrbeta insertional mutant.

The last three studies described here all demonstrate that the rate at which factor-independent mutants of a factor-dependent cell line are obtained can be increased over background by using retroviral vectors as insertional mutagens. The cell frequency at which factor-independent mutants were obtained was proportional to the number of vector integrants screened and one mutant was generated for every 10^7 integrants. The phenotype of factor-independent mutants differed between cell lines; with autocrine secretion of growth factors being a common mechanism in some cell lines but not in others. A detailed analysis of integration sites was not possible at the time many of these studies were performed, and only rarely could it be demonstrated that virus/vector integration was directly responsible for transformation.

Table 3.1. Summary of selected retroviral insertional mutagenesis studies in fact cell lines

Cell line	Description	Mutagen	Phenotype	Insertions	Frequency
D35	Mouse promyelocytic IL-3/GM-CSF dependent	WT Friend SFFV WT Rauscher MLV IAP	9/11 mutants produce autocrine factor (6/9 GM-CSF; 3/9 IL-3)	*GM-CSF producing mutants:* MLV/SFFV insertions 5' of GM-CSF gene in 3 mutants IAP insertion in GM-CSF 3'UTR in one mutant *IL-3 producing mutants:* IAP insertions 5' of IL-3 in 1-2 mutants	2.4×10^{-7} cells
			2/11 mutants have truncation of β_c extracellular domain plus deletion of β_{IL-3}	No vector insertion, but large chromosomal deletion. (Probably spontaneous mutants)	
FDC-Pmix	Mouse multipotent IL-3 dependent	IAP	3/7 IL-3 secretion	IAP Insertions 5' of IL-3	Mock (IAP): 3.6×10^{-9} cells
			4/7 No growth factor secretion		
FDC-P1(M)	Mouse promyelocytic IL-3 dependent	IAP	12/14 GM-CSF secretion 2/14 No growth factor secretion		Mock (IAP): 1.7×10^{-8} cells
TF-1	Human erythroleukaemia cell line GM-CSF dependent	MPSV RV vector	13% autocrine mechanism	No insertion sites analysed	Mock: 6.7×10^{-8} cells Vector: 1.5×10^{-6} cells 2.3×10^{-7} integrant
			85% other mechanisms	No insertion sites analysed 11/14 mutants analysed express constitutively	

Table 3.1. Summary of selected retroviral insertional mutagenesis studies in factor-dependent cell lines

Cell line	Description	Mutagen	Phenotype	Insertions	Frequency	References
FDC-P1(M)	Mouse promyelocytic IL-3 dependent	MPSV RV vector	Selection for GM-CSF α chain expressing mutants	No insertion sites analysed; rearrangements in GM-CSF α chain locus in 40% of mutants analysed; probably a lot are spontaneous mutants	Mock (IAP): 5.4×10^{-8} cells Vector: 3.0×10^{-7} cells 1.4×10^{-7} integrants	(Laker, Friel et al. 2000)
FDC-P1(M)	Mouse promyelocytic IL-3 dependent	MPSV RV vector	4/45 secreted IL-3 3/45 secreted GM-CSF 38/45 unknown other mechanisms	1 mutant contained RV insertion 997bp upstream of Pdgfrβ in reverse orientation	Mock (IAP): 4.6×10^{-8} cells Vector: 2.7×10^{-7} cells 1.5×10^{-7} integrants	(Meyer, Laker et al. 2002)

3.1.4.Retroviral vector insertional mutagenesis in the BAF3 cell line to identify anti-apoptotic genes

As mentioned, retroviral vectors derived from myeloproliferative sarcoma virus (MPSV) had previously been used to obtain IL-3 independent insertional mutants of the BAF3 cell line (Thomas *et al.* 1998; Mathieu *et al.* 2001). The main aim of these early studies was to identify genes involved in the regulation of apoptosis, therefore mutants that remained viable but did not proliferate in the absence of IL-3 were specifically selected for. Retroviral vector integration sites in IL-3 independent mutants were used as tags to identify these anti-apoptotic genes. Insertional activation of anti-apoptotic genes confers short-term resistance to IL-3 deprivation to these BAF3 mutants. BAF3 cells that overexpress Bcl-X_L survive but do not proliferate in the absence of IL-3 (Collins *et al.* 1992; Thomas *et al.* 1998). Upon IL-3 withdrawal they enter G_1 arrest, but lose viability after several days.

The selection protocol for apoptosis-resistant BAF3 mutants in these early studies was as follows (Thomas *et al.* 1998): BAF3 target cells were transduced on Day 1 and then cultured in the presence of IL-3 for the next 48 hours. IL-3 withdrawal was initiated on Day 3 for a total of approximately 60 hours. Firstly the cells were cultured in complete medium without IL-3 for 30 hours by which time ~95% of the cells had died. Dead cells were then eliminated by Ficoll density gradient; the remaining live cells were cloned in 1.5% Methocel MC semi-solid medium for another 30 hours in the absence of IL-3. IL-3 was then added back to these cultures. After 10 days individual clones were expanded in liquid medium containing IL-3 before characterisation for their ability to resist apoptosis upon IL-3 withdrawal. In these experiments, IL-3 independent mutants were obtained at a cell frequency of 1.4×10^{-7} and integrant frequency of 1.2×10^{-6}. Out of the 19 IL-3 independent mutants that were analysed, 4 (21%) were resistant to multiple apoptosis inducing pathways. These 4 mutants were found to have up-regulated Bcl-X_L expression due to retroviral vector / solo LTR insertion between 928 and 200bp upstream of the Bcl-X gene locus.

3.1.5.Pilot insertional mutagenesis experiment in the BAF3 cell line

It was recognised that the duration of IL-3 starvation impacts on the type of insertional mutant that is eventually selected for. As was discussed in section 3.1.2. on BAF3 transformation, insertional activation of 3 types of genes could result in IL-3 independence of BAF3 cells. These are genes encoding growth factors or their receptors, genes encoding intracellular signalling proteins and genes encoding (or regulating) anti-apoptotic proteins.

Mutants in which anti-apoptotic genes are activated by insertional mutagenesis can only be identified when IL-3 is added back to the medium. This might not however be necessary to identify mutants that have insertional activation of the 2 other types of genes.

In a pilot experiment (Pilot, Table 4.1.) performed before I joined the lab, it was attempted at first to select for mutants without adding IL-3 back to the medium. In this experiment 10^8 BAF3 cells were transduced with HV lentiviral vector (depicted in Figure 2.1.a.) at a MOI of 0.04. Selection for IL-3 independent clones was started on day 3 (48 hours post-transduction). 10^8 BAF3 cells were plated in 24 well plates at a cell density of 10^6 cells/ml. Cells were starved of IL-3 for 5 days, but no IL-3 independent clones grew out during this time. On day 8, cells in the mock and vector transduced arms of the experiment were rescued by adding IL-3 back to the medium, which surprisingly resulted in proliferating cells in all wells. On day 17, 8 randomly selected wells from each arm of the experiment were washed off IL-3. In the HV vector transduced cultures, all 8 wells contained GFP positive IL-3 independent clones. These clones were later shown to be one and the same mutant, named HV A2. Mutant HV A2 was able to proliferate long term in the absence of IL-3.

3.1.5.1. Important observations made in the pilot experiment

Several important observations towards the development of an insertional mutagenesis assay were made in this pilot experiment. Firstly, one long term IL-3 independent mutant (HV A2) was obtained at a relative low integrant frequency of 1 in 4 x 10^6, suggesting the HV lentiviral vector could transform BAF3 cells to IL-3 independence. Secondly, HV A2 was unable to expand at initial selection in complete medium without IL-3 and only grew out after rescue by adding IL-3 back to the culture. Thirdly, it was found that upon re-addition of IL-3 to the cultures, many mock-transduced parental cells were also rescued. Therefore, 5 day IL-3 starvation alone was clearly not sufficient to completely eliminate parental BAF3 cells.

3.1.6.Aims

To address these problems, there was firstly a need to explore treatments that in combination with IL-3 withdrawal could more effectively eliminate untransduced parental BAF3 cells without affecting the chance of obtaining IL-3 independent insertional mutants. X-ray irradiation in combination with IL-3 withdrawal was tested. The effect of this treatment on parental BAF3 and the, at the time uncharacterised, mutant HV A2 were investigated.

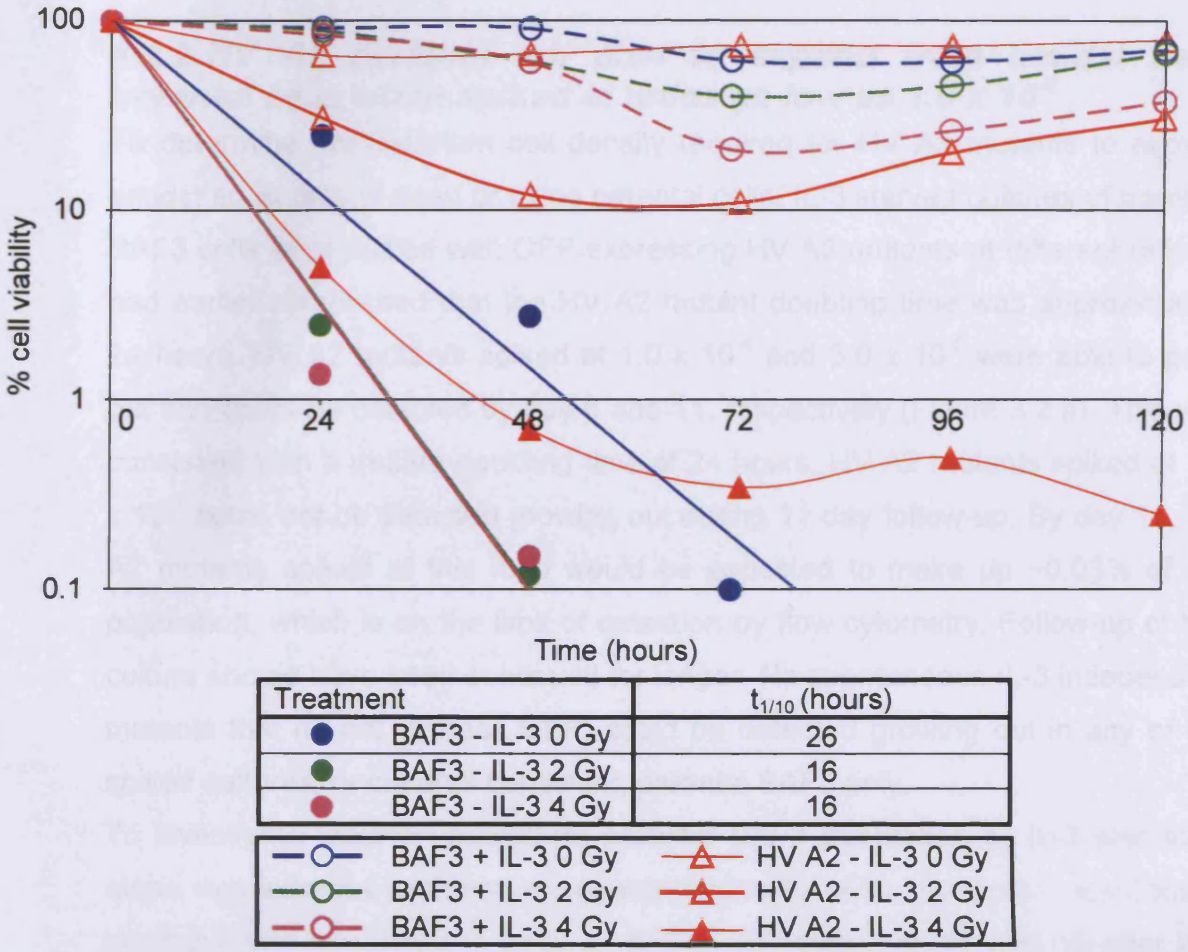
Secondly, there was a need to quantify how many rare mutants could grow out amidst an excess of dead or dying cells. For this reason, cultures in which parental BAF3 cells were spiked with HV A2 mutants at different ratios were set up and starved of IL-3. The outgrowth of HV A2 mutant cells in the cultures was monitored. The results of these experiments would enable us to revise our mutagenesis assay conditions.

3.2. Results

3.2.1.X-irradiation after IL-3 withdrawal accelerates the rate at which parental BAF3 cells undergo apoptosis, but can still be used to select IL-3 independent mutants

Treatments were explored that, in combination with IL-3 withdrawal, would allow the selective elimination of parental BAF3 cells without affecting mutant HV A2 cell viability. Serum starvation was tested first, but this treatment did not differentially affect BAF3 and mutant HV A2 cell viability with both cells dying at similar rates (Data not shown).

BAF3 cells cultured in the presence of IL-3 are largely resistant to the effects of X-irradiation induced DNA damage. However, the rate at which BAF3 cells lose viability can be increased if cells are X-irradiated shortly after IL-3 withdrawal (Collins *et al.* 1992). The ability of the HV A2 mutant to resist X-irradiation had not yet been tested. Consistent with what was reported previously, BAF3 cultured in the presence of IL-3 survived X-irradiation at the two doses tested (Figure 3.1.). A decrease in cell viability could be seen between 36 and 72 hours post-irradiation; this was most evident at the higher dose of 4 Gy. From 72 hours onwards, cell viability increased again indicating that those BAF3 cells that had survived X-irradiation resumed proliferation. In the absence of IL-3, X-irradiation at 2 and 4 Gy accelerates the rate at which parental BAF3 cells undergo apoptosis. The HV A2 mutant is somewhat protected from X-irradiation induced apoptosis. At 4 Gy, HV A2 cell viability decreased to less than 1% within 48 hours of irradiation, but unlike parental BAF3 irradiated at the same dose some viable cells still remain detectable by flow cytometry over the next 3 days. At 2 Gy, the rate at which HV A2 lose viability is markedly slower and after 72 hours the remaining viable cells resume proliferation.



Treatment	$t_{1/10}$ (hours)
BAF3 - IL-3 0 Gy	26
BAF3 - IL-3 2 Gy	16
BAF3 - IL-3 4 Gy	16

BAF3 + IL-3 0 Gy	HV A2 - IL-3 0 Gy
BAF3 + IL-3 2 Gy	HV A2 - IL-3 2 Gy
BAF3 + IL-3 4 Gy	HV A2 - IL-3 4 Gy

Figure 3.1. X-irradiation after IL-3 withdrawal accelerates the rate at which parental BAF3 cells undergo apoptosis, but can still be used to select IL-3 independent mutants

5 x 10^6 BAF3 (circles) and HV A2 (triangles) cells were washed thrice in complete medium without IL-3 and then X-irradiated with 0, 2 or 4 Gy. Following irradiation, cells were cultured in the presence (dashed lines) or absence (full lines) of IL-3. Cell viability in these cultures was measured by propidium iodide (PI) exclusion every day for 5 days.

BAF3 were cultured in the presence of IL-3 at a cell density of 2 x 10^5 cells/ml. There is a transient decrease in cell viability between 36 and 72 hours post-irradiation which is most pronounced in cultures irradiated with 4 Gy. Cell viability increases from 72 hours post-irradiation, indicating that some BAF3 survived irradiation in the presence of IL-3 and resumed proliferation.

BAF3 were cultured in the absence of IL-3 at a cell density of 10^6 cells/ml. Exponential trendlines were fitted to the BAF3 0 Gy, BAF3 2 Gy and BAF3 4 Gy data series. R-squared values were 0.9505, 0.9959 and 0.9599, respectively. The time at which 10% ($t_{1/10}$) of the original number of cells are still viable after each treatment is shown in the figure legend. It was calculated using the formula $y = c * e^{rx}$, where c is the intercept (set at 100), e is the base of the natural logarithm and r is the decay constant. The decay constants for the BAF3 0 Gy, BAF3 2 Gy and BAF3 4 Gy data series were -0.0879, -0.143 and -0.1442, respectively. The rate at which BAF3 cultured in the absence of IL-3 undergo apoptosis is accelerated after irradiation.

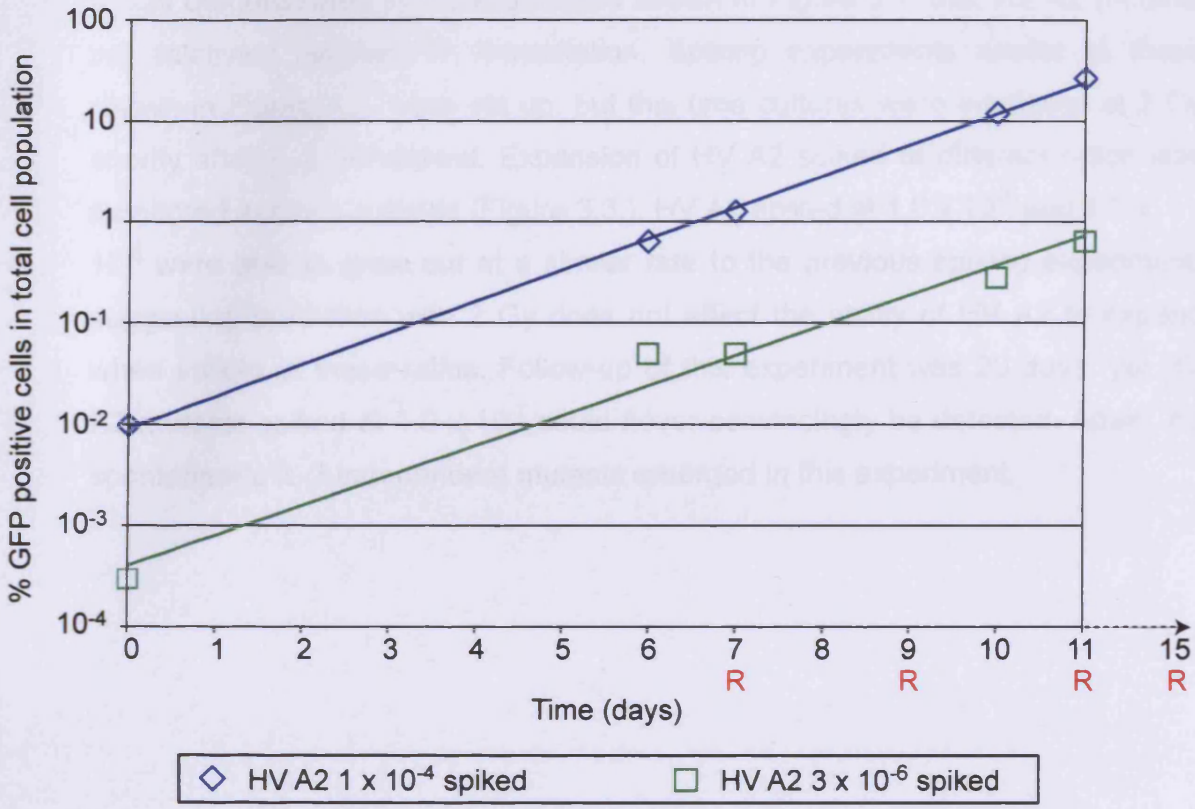
HV A2 were cultured in the absence of IL-3 at a cell density of 2 x 10^5 cells/ml. X-irradiation with 2 Gy leads to a 72 hours transient decrease in HV A2 cell viability. HV A2 irradiated at 4 Gy lose cell viability at a similar rate to parental BAF3. X-irradiation with 2 Gy therefore seems to be a suitable treatment to more effectively eliminate parental BAF3 cells, without affecting the ability to select mutants.

3.2.2.HV A2 mutants are able to expand from amidst dead parental cells when spiked at ratios as low as 1.0 x 10^{-6}

To determine the minimum cell density required for HV A2 mutants to expand amidst an excess of dead or dying parental cells, IL-3 starved cultures of parental BAF3 cells were spiked with GFP expressing HV A2 mutants at different ratios. I had earlier established that the HV A2 mutant doubling time was approximately 24 hours. HV A2 mutants spiked at 1.0 x 10^{-4} and 3.0 x 10^{-6} were able to grow out and could be detected by day 6 and 11, respectively (Figure 3.2.a). This was consistent with a mutant doubling time of 24 hours. HV A2 mutants spiked at 1.0 x 10^{-7} could not be detected growing out during 11 day follow-up. By day 11, HV A2 mutants spiked at this ratio would be expected to make up ~0.03% of the population, which is on the limit of detection by flow cytometry. Follow-up of this culture should have been continued for longer. No spontaneous IL-3 independent mutants that do not express GFP, could be detected growing out in any of the spiked cultures, or cultures containing parental BAF3 only.

To investigate whether incomplete parental BAF3 elimination by IL-3 starvation alone was indeed a problem, as experienced in the pilot experiment, fractions of each culture were rescued by IL-3 addition on days 7, 9, 11 and 15 after IL-3 withdrawal. In all cultures, including the one containing BAF3 only, cells could be effectively rescued by IL-3 at all time points. The percentage of GFP positive HV A2 mutants in the rescued cultures was measured (Figure 3.2.b). In cultures spiked at 1.0 x 10^{-4} and 3.0 x 10^{-6} and rescued on day 7 and 9, HV A2 mutants could be fully selected, as indicated by the high (>90%) percentage of GFP positive cells in these cultures. HV A2 spiked at 1.0 x 10^{-7} could not be fully selected upon rescue with IL-3.

a.



b.

% GFP positive live cells in cultures rescued with IL-3 on days 7, 9, 11 and 15

Rescue	1.0×10^{-4}	3.0×10^{-6}	1.0×10^{-7}	BAF3 only
Day 7	97.60	93.52	0.00	0.00
Day 9	89.52	94.20	0.62	0.00
Day 11	ND	ND	0.56	0.00
Day 15	ND	ND	0.09	0.00

Figure 3.2. IL-3 independent HV A2 mutants are able to grow out in the presence of dead parental BAF3 cells when spiked at ratios of 1.0×10^{-4} and 3.0×10^{-6}.

5×10^7 parental BAF3 cells were washed thrice in complete medium without IL-3 and spiked with HV A2 mutants at ratios of 1.0×10^{-4}, 3.0×10^{-6} and 1.0×10^{-7}. The negative control contained parental BAF3 cells only. Cells were cultured in the absence of IL-3 at a density of 10^6 cells/ml.

a. Cell viability in these cultures was measured by propidium iodide (PI) exclusion on days 6, 7, 10 and 11. A GFP positive live cell population could be detected in cultures spiked 1.0×10^{-4} and 3.0×10^{-6} with HV A2 cells by days 6 and 11, respectively. No HV A2 cells spiked at 1.0×10^{-7} grew out during 11 day follow-up. In the unspiked BAF3 culture no spontaneous IL-3 independent mutants emerged.

b. Fractions of each of the four cultures were rescued by IL-3 re-addition on days 7, 9, 11 and 15, as indicated by "R" on the chart. Growing cells emerged in all four cultures rescued at each time point, including in the unspiked cultures containing only parental BAF3 cells. The percentage of GFP positive live cells in these rescued cultures was measured by flow cytometry two weeks after rescue and is shown in the table. In cultures spiked 1.0×10^{-4} and 3.0×10^{-6} rescued on days 7 and 9, 90% or more of the cells were GFP positive. In cultures spiked 1.0×10^{-7}, HV A2 mutants could not be fully selected after rescue with IL-3. (ND = not done)

It was demonstrated in the experiment shown in Figure 3.1. that HV A2 mutants are relatively resistant to X-irradiation. Spiking experiments similar to those shown in Figure 3.2. were set up, but this time cultures were irradiated at 2 Gy shortly after IL-3 withdrawal. Expansion of HV A2 spiked at different ratios was monitored in these cultures (Figure 3.3.). HV A2 spiked at 1.0×10^{-5} and 1.0×10^{-6} were able to grow out at a similar rate to the previous spiking experiment, suggesting irradiation with 2 Gy does not affect the ability of HV A2 to expand when spiked at these ratios. Follow-up of this experiment was 20 days, yet HV A2 mutants spiked at 1.0×10^{-7} could never convincingly be detected. Again, no spontaneous IL-3 independent mutants emerged in this experiment.

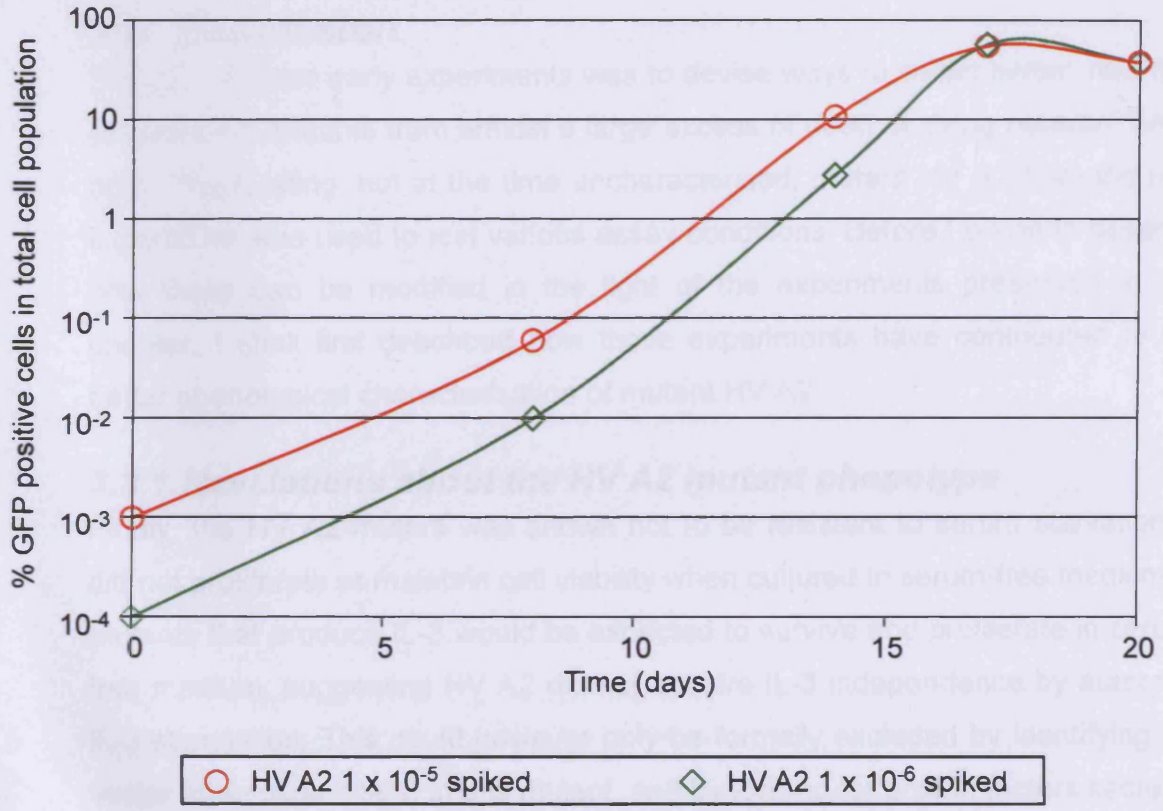


Figure 3.3. Following irradiation with 2 Gy, HV A2 cells are able to grow out in the presence of dead parental BAF3 cells when spiked at ratios of 1.0×10^{-5} and 1.0×10^{-6}.

Similar to the experiment shown in Figure 3.2., 5×10^7 parental BAF3 cells were washed thrice in complete medium without IL-3 and spiked with HV A2 mutants at ratios of 1.0×10^{-5}, 1.0×10^{-6} and 1.0×10^{-7}. The negative control contained parental BAF3 cells only. Cells were then X-irradiated with 2 Gy and cultured in the absence of IL-3 at a density of 10^6 cells/ml.
Cell viability in these cultures was measured by propidium iodide (PI) exclusion on days 8, 14, 17 and 20. A GFP positive live cell population could be detected in cultures spiked with 1.0×10^{-5} and 1.0×10^{-6} HV A2 cells after two weeks. No HV A2 cells spiked at 1.0×10^{-7} grew out during 20 day follow-up. In the unspiked BAF3 cultures no spontaneous IL-3 independent mutants emerged.

3.3. Discussion

The aim of these early experiments was to devise ways to select better, rare IL-3 independent mutants from amidst a large excess of dead or dying parental BAF3 cells. The existing, but at the time uncharacterised, mutant HV A2 from the pilot experiment was used to test various assay conditions. Before I go on to describe how these can be modified in the light of the experiments presented in this chapter, I shall first described how these experiments have contributed to the better phenotypical characterisation of mutant HV A2.

3.3.1. Revelations about the HV A2 mutant phenotype

Firstly, the HV A2 mutant was shown not to be resistant to serum starvation. It did not proliferate or maintain cell viability when cultured in serum-free medium. Mutants that produce IL-3 would be expected to survive and proliferate in serum-free medium, suggesting HV A2 did not acquire IL-3 independence by autocrine IL-3 stimulation. This could however only be formally excluded by identifying the vector integration site(s) in this mutant, and by looking for growth factors secreted in the culture medium by these cells. Serum independent growth was further reported of BAF3 transformed with the oncogenic human tyrosine kinase Brk (Kasprzycka *et al.* 2006). Serum-independence is a very robust phenotype only a very limited number of mutants are likely to display, hence it is not a good method to selectively eliminate parental cells from mutants.

Secondly, HV A2 cells were shown to be resistant to X-irradiation with up to 2 Gy. Again, autocrine IL-3 producing mutants would expected to be resistant to X-irradiation, similar to BAF3 cells cultured in the presence of IL-3. It was shown that X-irradiation induces expression of the pro-apoptotic cell surface receptor Fas on BAF3 cells (Gutierrez del Arroyo *et al.* 2000). X-irradiation with 4 Gy leads to single stranded DNA breaks. This damage is detected by ATM and ATR kinases that phosphorylate and thereby activate p53. X-irradiation was found to increase p53 levels in BAF3 cells (Palacios *et al.* 2000). Fas is constitutively expressed at low levels on BAF3 cells; a p53-mediated increase in its cell surface expression sensitises BAF3 cells to extracellular death ligands, such as Fas

ligand, and could account for the decrease in cell viability observed following X-irradiation. Fas cell surface expression is not induced by IL-3 withdrawal per se, but its up-regulation is accelerated in X-irradiated cells which are deprived of IL-3. IL-3 can rescue BAF3 cells from X-irradiation induced apoptosis by reducing Fas receptor levels on the cell surface. Resistance to X-irradiation induced apoptosis was also correlated with Bcl-X_L expression in BAF3 cells (Mathieu *et al.* 2001). Therefore, the mechanism responsible for IL-3 independence of HV A2 cells is likely to induce some degree of Bcl-X_L expression, but maybe to lower levels than that induced by IL-3 in parental BAF3 cells, explaining the partial resistance to X-irradiation of HV A2.

3.3.2. More effective elimination of parental BAF3 cells after X-irradiation

Parental BAF3 cell death upon IL-3 withdrawal was confirmed to be asynchronous, as demonstrated by the very effective rescue of parental BAF3 cells by IL-3 even after two weeks of IL-3 starvation. Asynchronous parental BAF3 cell death is not a problem if IL-3 independent clones were able to expand without the need for rescue with IL-3, unfortunately this is not the case. Irradiation can accelerate decay of parental cells upon IL-3 withdrawal. There is however a possibility that a very small but still significant number of cells may survive X-irradiation and can be recovered upon rescue with IL-3. Rescue with IL-3 of irradiated spiked cultures was not tested. If IL-3 rescue is required for mutants to grow out; irradiation might reduce background rescue of parental cells and mutants of a phenotype similar to HV A2 should be able to survive at low doses of irradiation. Upon rescue, IL-3 can then stimulate the proliferation of these residual viable mutants.

3.3.3.HV A2 mutants can expand amidst dying parental BAF3 cells when spiked at ratios as low as 10^{-6}

In the pilot experiment, outgrowth of IL-3 independent mutants was only possible after rescue with IL-3. The series of spiking experiments go some way in explaining the need for rescue with IL-3. Only mutants spiked at ratios of 10^{-6} or higher were able to grow out. The mutant from the pilot experiment was obtained at a cell frequency of 10^{-8}, so not surprisingly this mutant only grew out after rescue.

To have some idea as to how many integrants would need to be screened in mutagenesis experiments in order for mutants to grow out, quantitative considerations should be made. The distance from the integration site at which retroviruses are able to interfere with neighbouring gene expression is generally considered to be 10^3-10^4 base pairs (Baum *et al.* 2003; Baum *et al.* 2004). This means that theoretically 10^5 to 10^6 hits in a mouse genome of 2.7 x 10^9 base pairs could result in a mutant at a single locus. Going by the lower integrant frequency of 10^{-6}, 10^8 target cells would need to be transduced at MOI 1 to obtain 100 mutants at a single locus.

A predicted integrant frequency of 10^{-6} is consistent with that reported in earlier retroviral mutagenesis experiments in the BAF3 target cell. IL-3 independent mutants were obtained with an MPSV retroviral vector at a cell frequency of 1.4 x 10^{-7} and an integrant frequency of 1.2 x 10^{-6} (Thomas *et al.* 1998). However, in this paper selection in the absence of IL-3 was of much shorter duration than in our pilot experiment. An IL-3 independent clone was defined as "a clone where the percentage of viable cells after 20 hours of IL-3 starvation was at least 8 fold higher than the parental cells". Most of these clones were probably not long term IL-3 independent and also it was not established whether they were all independent clones. Hence, the frequency figures quoted above are likely to be overestimates. The integrant frequency in particular seems to be very high; as in studies using other factor dependent cell lines this was in the order of 1.4 – 2.3 x 10^{-7}, 10 fold lower (Stocking *et al.* 1993; Laker *et al.* 2000; Meyer *et al.* 2002)(See Table 3.1.). In the pilot experiment with the HV lentiviral vector in

BAF3 cells, the HV A2 mutant was obtained at an integrant frequency of 2.5 x 10^{-7}.

From the spiking experiments we now know that to obtain IL-3 independent mutants without rescue we need at least 100 mutants per 10^8 target cells. If the HV integrant frequency is really 2.5 x 10^{-7}, that would mean 10^8 target cells would have to be transduced with the HV lentiviral vector at MOI 4 at least, for mutants to grow out without rescue. As a compromise between a predicted integrant frequency of 10^{-6} and an observed integrant frequency of 2.5 x 10^{-7} in the pilot experiment, it was decided to transduce 10^8 target cells at MOI 2 with the HV lentiviral vector in mutagenesis experiments that will be described in the next chapter.

Chapter 4

4. Insertional mutagenesis by a lentiviral vector

4.1. Introduction

The pilot experiment described in section 3.1.5. demonstrated that the HV lentiviral vector could transform BAF3 cells to IL-3 independence. The HV lentiviral vector, depicted in Figure 2.1.a, has two wild-type HIV-1 LTRs and an internal SFFV LTR promoter driving GFP expression (Demaison *et al.* 2002; Ikeda *et al.* 2003).

In this chapter, I used the HV vector to select for IL-3 independent insertional mutants of the BAF3 cell line. At first the modified assay conditions informed by the results in chapter 3 were tested. Target cells were irradiated prior to selection to reduce background survival of parental BAF3 cells. I aimed to screen a sufficient number of integrants to enable mutants to expand in the presence of an excess of dying cells without the need for rescue with IL-3. Assay conditions were further modified to allow more efficient and reproducible mutant detection. This involved switching to a BAF3 cell derivative which over-expresses human Bcl2. I aimed to establish a frequency at which IL-3 independent mutants of this cell line could be obtained after transduction with the HV lentiviral vector. Any IL-3 independent mutants obtained with this vector were characterised in detail. In the previous chapter about half a dozen potential target genes, such as IL-3, IGF-1, Ghr, Pdgfr, Fgfr1 and Bcl-X, that could transform BAF3/BCL15 cells to IL-3 independence were identified. Vector integration site analysis of mutants revealed which of these genes was responsible for transformation. The mechanism whereby the vector affects expression of this target gene was investigated.

4.2. Results

4.2.1.HV lentiviral vector mutagenesis experiments in the BAF3 cell line

In mutagenesis experiments in the BAF3 cell line, the number of integrants analysed was determined by GFP expression. Almost 5 x 10^8 BAF3 cells were transduced with the HV lentiviral vector in 6 independent mutagenesis experiments (Table 4.1.), screening approximately 8 x 10^8 HV vector integrants in the process. 10^8 target cells each were transduced at MOI 1 and 0.7, respectively in experiments LV1 A and B, in a single round of transduction. According to the calculations described in the discussion to chapter 3, ideally this number of target cells would need to be transduced at MOI 2 to pick up ~100 mutants. In experiment LV1 B the vector preparation needed to achieve high MOI proved to be highly toxic to target cells with half of them dying. It was therefore decided in subsequent experiments (LV1 C to F) to subject target cells to serial rounds of transduction to achieve high MOI.

In experiments LV1 A and B, IL-3 was withdrawn from cultures on Day 3, at 48 hours after transduction, and this was followed immediately by irradiation with 2 Gy. 10^7 cells were cultured in 10 ml medium without IL-3 in T25 tissue culture flasks. Cultures were fed once a week with fresh medium. Samples of these cultures were analysed by flow cytometry on a biweekly basis to monitor outgrowth of GFP positive, PI excluding populations i.e. possible mutants. It was hoped mutants could be detected without the need for rescue with IL-3. However, in case this should be required, irradiation would be expected to have efficiently eliminated any parental BAF3 and mutants would be selectively rescued. Half of the culture in each T25 flask from LV1 A and B was rescued by adding IL-3 back to the medium on days 30 and 14 of the experiment, respectively. Neither mutant nor parental cells could be rescued at either time point. Unrescued cultures were followed up for 6 weeks, but no mutants could be detected growing out. Irradiation with 2 Gy may have been too harsh a treatment; it was therefore decided to abandon this approach in subsequent experiments.

Table 4.1. HV lentiviral vector mutagenesis frequencies

Experiment	Target Cell[d]	Mock Mutants		HV Vector Transduced Mutants		
		Mutant Number	Cell Frequency [e]	Mutant Number	Cell Frequency [e]	Integrant Frequency [f]
Pilot	BAF3	0	$< 1.0 \times 10^{-8}$	1	1.0×10^{-8}	2.5×10^{-6}
LV1 A [a, b]	BAF3	0	$< 1.1 \times 10^{-8}$	0	$< 1.2 \times 10^{-8}$	$< 1.2 \times 10^{-8}$
LV1 B [a, b]	BAF3	0	$< 1.3 \times 10^{-8}$	0	$< 2.1 \times 10^{-8}$	$< 3.1 \times 10^{-8}$
LV1 C [b]	BAF3	0	$< 1.4 \times 10^{-8}$	0	$< 1.2 \times 10^{-8}$	$< 3.4 \times 10^{-9}$
LV1 D [b]	BAF3	0	$< 1.2 \times 10^{-8}$	0	$< 1.3 \times 10^{-8}$	$< 5.0 \times 10^{-9}$
LV1 E [b]	VI34	0	$< 1.1 \times 10^{-8}$	0	$< 1.1 \times 10^{-8}$	$< 4.8 \times 10^{-9}$
LV1 F	BAF3	0	$< 1.2 \times 10^{-8}$	0	$< 1.3 \times 10^{-8}$	$< 5.1 \times 10^{-9}$
LV2 [c, g]	BCL15	0	$< 1.0 \times 10^{-7}$	2	2.0×10^{-7}	5.0×10^{-8} 1.7×10^{-8}
LV3 [c, g]	BCL15	2	4.0×10^{-8}	13	2.6×10^{-7}	1.8×10^{-7} 6.8×10^{-8}

Table 4.1. HV lentiviral vector mutagenesis frequencies

1-10 x 10^7 IL-3 dependent target cells were transduced with the HV lentiviral vector. The cell and integrant frequencies at which IL-3 independent mutants were obtained following transduction with this vector are shown in the table.

[a] Target cells in experiments LV1 A and B were irradiated with 2Gy immediately after IL-3 withdrawal on day 3 of the experiment.
[b] Selection for IL-3 independent mutants in experiments LV1 A to E was performed in T25 tissue culture flasks in which cells were cultured at a density of 10^6 cells/ml in 10ml volume. In all other experiments, including the pilot experiment, selection for IL-3 independent mutants was performed in 24 well plates at cell densities ranging from 2.5-10 x 10^5 cells/ml.
[c] Selection for IL-3 independent mutants in experiments LV2 and 3 was performed according to the "standard selection protocol". Cells were washed of IL-3 on day 3/4 and plated in 24 well plates at 5 x 10^5 cells per well in 2ml complete medium without IL-3. IL-3 was added back to the cultures on day 10/11 by replacing 1ml medium with complete medium containing IL-3. Cultures reached confluence after 2-3 days and cells were again washed of IL-3. They were then cultured in T25 flasks in 10ml medium without IL-3.
[d] BAF3 derivative cell lines VI34 and BCL15 over-express Bcl-X_L and Bcl-2, respectively.
[e] The cell frequency was calculated by dividing the number of mutants by the total number of target cells in each experiment. Vector associated toxicity to cells was seen in some of the LV1 experiments. The number of target cells was therefore adjusted for the number of live cells analysed in each arm (mock or HV-transduced) of these experiments.
[f] The integrant frequency was calculated by dividing the number of mutants by the total number of integrants in each experiment. The number of "GFP" integrants was calculated by multiplying the number of live target cells by the MOI. The MOI was determined by flow cytometry for GFP expression at 48h to 72h post-transduction. The GFP integrant frequency is shown in green. In experiments LV2 and LV3, the average number of vector copies per cell was also determined by Q-PCR. This number was multiplied by the number of target cells to obtain the number of "Q-PCR" integrants. The Q-PCR integrant frequency is shown in black.
[g] The average cell frequencies at which IL-3 independent mutants were obtained in experiments LV2 and LV3 combined were calculated. The average mock cell frequency was 2 x 10^{-8}; the average HV cell frequency was 2.3 x 10^{-7} ($p < 0.001$; $X^2 = 176.4$). The average integrant frequencies were: 1.2 x 10^{-7} (GFP) and 4.3 x 10^{-8} (Q-PCR).

In experiment LV1 C and D, higher MOIs of 3.5 and 2.5, respectively, were achieved after 3 to 4 rounds of transduction of 4 hours each. After IL-3 withdrawal, 10^7 cells were again cultured in T25 flasks in 10ml volume. Half the culture in each flask in experiments LV1 C and D were rescued with IL-3 on days 20 and 14, respectively. Again, neither mutants nor parental cells could be rescued at either time point. This was surprising as cells cultured under similar conditions in the experiment shown in Figure 3.3. were effectively rescued by IL-3 after having been grown in the absence of IL-3 for up to two weeks. Cells in this experiment were cultured at the same cell density as those in LV1 C and D. BAF3 cells in the pilot experiment were also effectively rescued by IL-3, but were grown at a lower cell density of 2.5×10^5 cells per ml in 24 well plates.

In experiment LV1 F, I investigated whether cell density played a role in both the ability of mutants to grow out and the effectiveness of rescue with IL-3. Target cells were transduced at MOI 2.5 in a single round of transduction using a high titre HV vector preparation (5×10^7 BAF3 i.u./ml (by GFP expression), a titration curve of this vector batch is shown in Figure 2.2.). At the start of IL-3 withdrawal on day 3, a third each of the cells were cultured at 2.5×10^5, 5×10^5 and 10^6 cells/ml in 24 well plates. One week later on day 10, half of the wells in each of the three groups were rescued with IL-3. Cells in ~80% of these wells were rescued, both in the mock and HV transduced arms of the experiment. Upon reaching confluence, the cells in these wells were washed and starved of IL-3 again. In the neither the unrescued nor the rescued-re-starved cultures did any IL-3 independent mutants emerge.

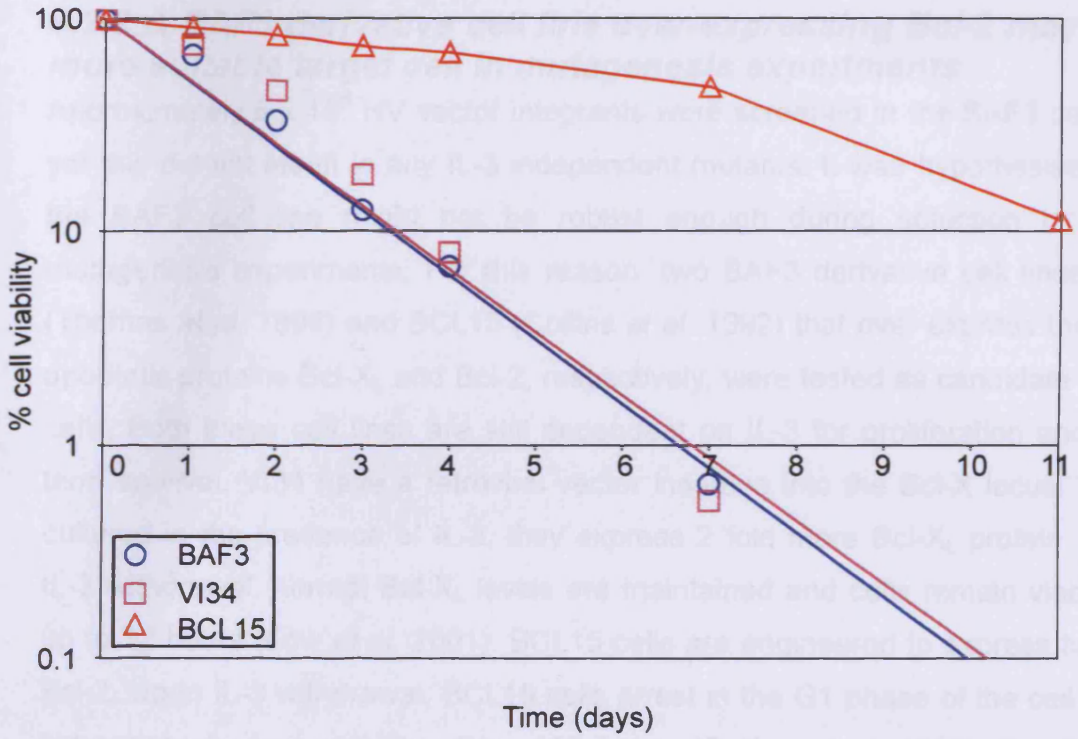


Figure 4.1. Survival of BAF3 cell line derivatives after IL-3 withdrawal.

Two BAF3-derivative cell lines VI34 (Thomas et al 1998) and BCL15 (Collins et al 1992), that over-express the anti-apoptotic proteins Bcl-X_L and Bcl-2 respectively, were tested for their ability to resist apoptosis upon IL-3 withdrawal in comparison to parental BAF3 cells.
Cells were washed thrice in complete medium without IL-3. Cell viability in these cultures was measured by propidium iodide (PI) exclusion on days 1, 2, 3, 4, 7 and 11.
Bcl-X_L over-expressing VI34 cells lose cell viability at a similar rate to BAF3 cells. Exponential trendlines were fitted to the BAF3 and VI34 data series. R-squared values were 0.9876 and 0.9511, respectively. The $t_{1/10}$ for both is approximately 80 hours, calculated as described in Figure 3.1. Bcl-2 over-expressing BCL15 cells maintain cell viability for up to 4 days following IL-3 withdrawal. Thereafter they lose cell viability but at a much slower rate than BAF3/VI34 cells.

4.2.2.A BAF3 derivative cell line over-expressing Bcl-2 may be a more suitable target cell in mutagenesis experiments

Approximately 8 x 10^8 HV vector integrants were screened in the BAF3 cell line, yet this did not result in any IL-3 independent mutants. It was hypothesised that the BAF3 cell line might not be robust enough during selection in these mutagenesis experiments. For this reason, two BAF3 derivative cell lines VI34 (Thomas *et al.* 1998) and BCL15 (Collins *et al.* 1992) that over-express the anti-apoptotic proteins Bcl-X_L and Bcl-2, respectively, were tested as candidate target cells. Both these cell lines are still dependent on IL-3 for proliferation and long term survival. VI34 have a retroviral vector insertion into the Bcl-X locus. When cultured in the presence of IL-3, they express 2 fold more Bcl-X_L protein. Upon IL-3 withdrawal, normal Bcl-X_L levels are maintained and cells remain viable for up to 32 hours (Low *et al.* 2001). BCL15 cells are engineered to express human Bcl-2. Upon IL-3 withdrawal, BCL15 cells arrest in the G1 phase of the cell cycle but eventually lose viability after ~100 hours (Collins *et al.* 1992). For BCL15 cells, these earlier observations were confirmed in Figure 4.1. However, the phenotype of VI34 cells upon IL-3 withdrawal proved to be less robust than reported. VI34 cells die at a similar rate to parental BAF3. VI34 were the target cell in experiment LV1 E. However, no mutants were obtained screening 2 x 10^8 integrants.

4.2.3.HV lentiviral vector mutagenesis in the Bcl-2 over-expressing BCL15 cell line.

5 x 10^7 BCL15 cells were transduced with HV lentiviral vector in experiment LV2 (Table 4.1). The number of vector integrants in all experiments with BCL15 cells was determined by quantitative PCR (Q-PCR) as well as by measuring GFP expression at 72 hours post-transduction. The number of Q-PCR integrants was usually 3-5 fold higher than the number of GFP integrants. The integrant frequencies quoted in the ensuing text were calculated using the number of Q-PCR integrants.

Cells in LV2 were selected from day 4 in 24 well plates at 5 x 10^5 cells per well in 2ml medium. One fifth of the wells were rescued with IL-3 on day 11, after one

week of IL-3 starvation. According to figure 4.1., 90% of BCL15 cells will have died at this time. However, in all wells in the mock and HV-transduced cultures to which IL-3 was added back, cells were effectively rescued. These cells were washed of IL-3 a second time round and cultured in T25 flasks in 8ml volume. Two weeks later, two IL-3 independent clones HV3 and HV14 (Figure 4.2.) grew out at a cell frequency of 2.0×10^{-7} and integrant frequency of 1.7×10^{-8} (Table 4.1.). This method of selection would become known as the standard selection protocol. No mutants emerged in wells not rescued by IL-3.

In experiment LV3, an additional 13 IL-3 independent mutants (HV10 /24 /25 /26 /32 /41 /42 /43 /48 /49 /72 /77 /82, Figure 4.2.) were obtained using the standard selection protocol, at a similar cell frequency to LV2 of 2.6×10^{-7} and an integrant frequency of 6.8×10^{-8} (Table 4.1). In this experiment, two mock mutants (M8 and M53) were also obtained at a cell frequency of 4.0×10^{-8}.

4.2.4. Characterisation of HV lentiviral vector insertional mutants

4.2.4.1. Clonality and vector copy number determination in IL-3 independent mutants

To determine how many unique IL-3 independent mutants were picked up in each experiment, Southern blots were performed (Figure 4.2.). Mutants' genomic DNA was digested with a restriction enzyme that cuts only once inside the vector. Blots were hybridised with the cDNA for GFP. Examination of the band pattern revealed no similarities between different mutants, suggesting they are all unique clones. The presence of different intensity bands within a lane would suggest the cell population is not entirely clonal and bystander cells may be present. Again, this was not observed. (It is common for smaller bands to show up with greater intensity, because these DNA fragments are more efficiently transferred from gel to membrane). HV vector copy number could also be determined in each of the mutants. The pilot experiment was performed at low MOI of 0.04 (GFP); mutant HV A2 contains just a single copy of the vector. Experiment LV2 was carried out at MOI 12 (Q-PCR); the vector copy number in mutants HV3 and HV14 is 12 and 7, respectively. Experiment LV3 was performed at MOI 3.8 (Q-PCR); LV3 mutants have 1 – 4 copies of the vector per genome.

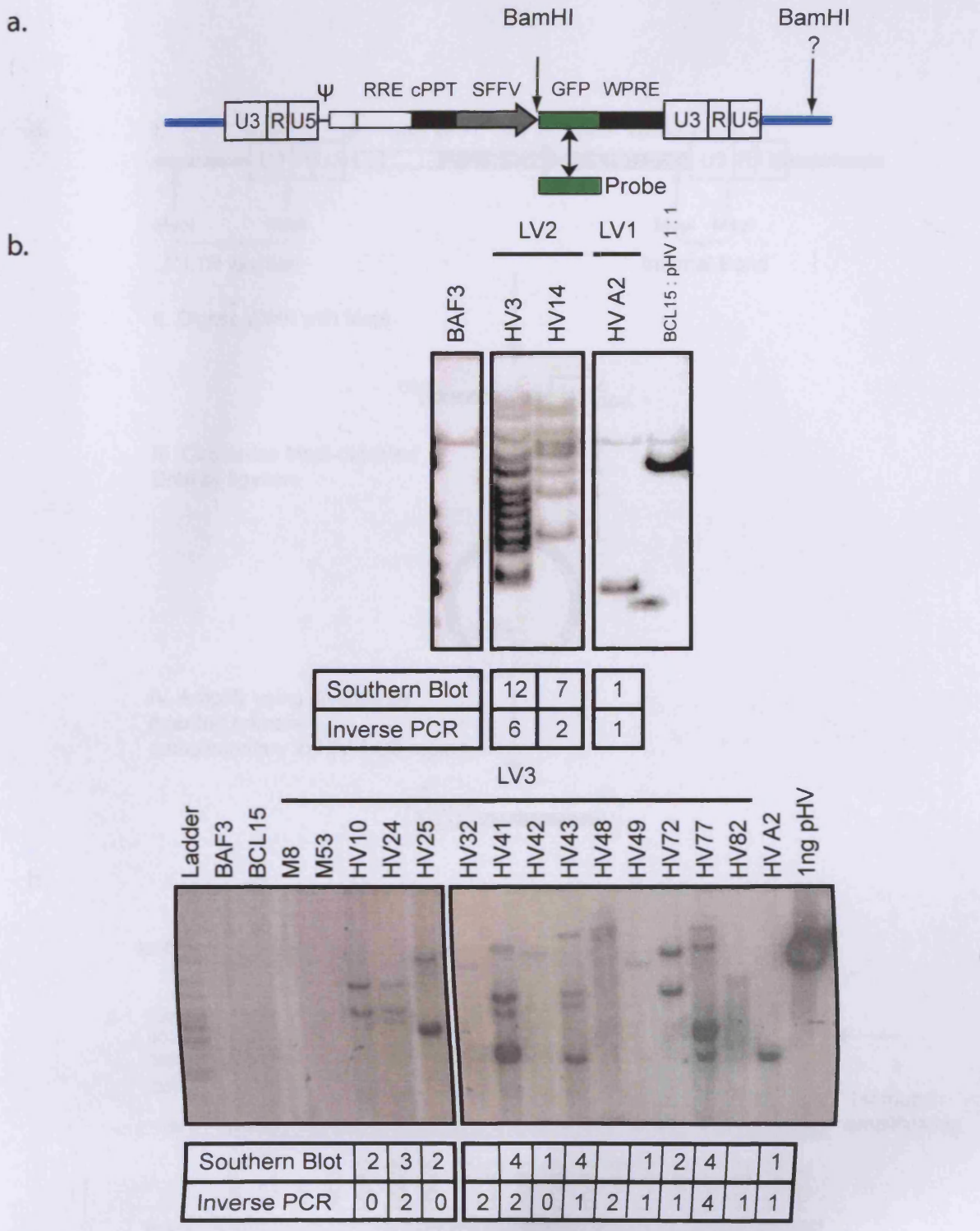


Figure 4.2. Southern blots of IL-3 independent mutants obtained in experiments LV1 – 3.

a. 15µg genomic DNA of each mutant was digested with BamHI, which cuts only once inside the vector. Blots were probed with the cDNA for eGFP. Band patterns were examined to determine clonality of the mutants and HV vector copy number per genome. b. A phosphorscreen image is shown of blots from LV1 and 2. Blots from LV3 were exposed to autoradiography film. Irrelevant lanes on blots were removed. The tables below each blot show the number of bands seen on the Southern blot as well as the number of integration sites for each mutant that were cloned out by inverse PCR. (Mutant HV26 is missing from this figure.)

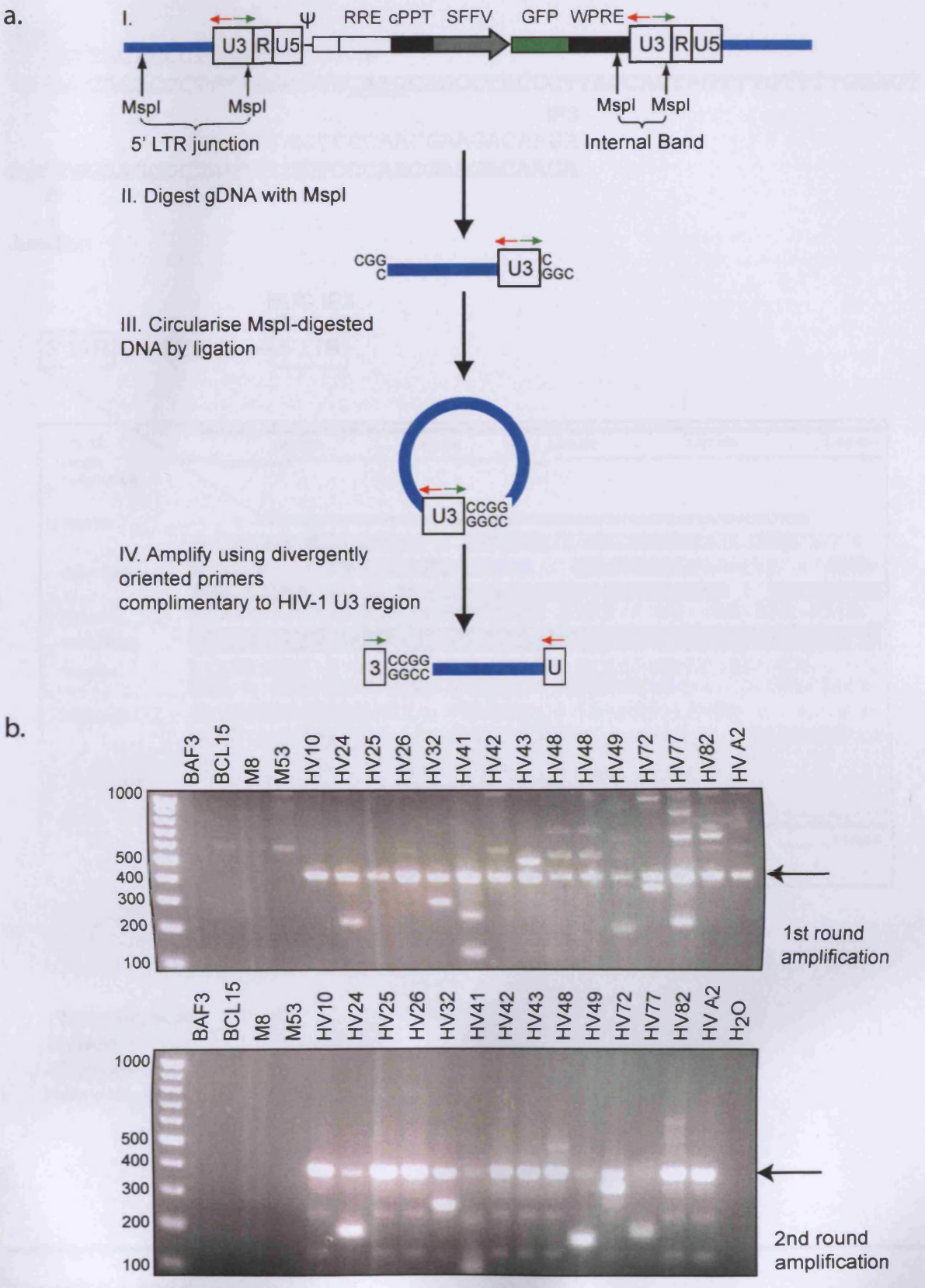


Figure 4.3. a. and b.

c.

HUC

CTTGTTACACCCTGTGAGCCT →

TTTGTTACACCCTGTGAGCCTGCATGCAGCCTACCCTTAGCAGTAGTTTGTTTTGGAGT

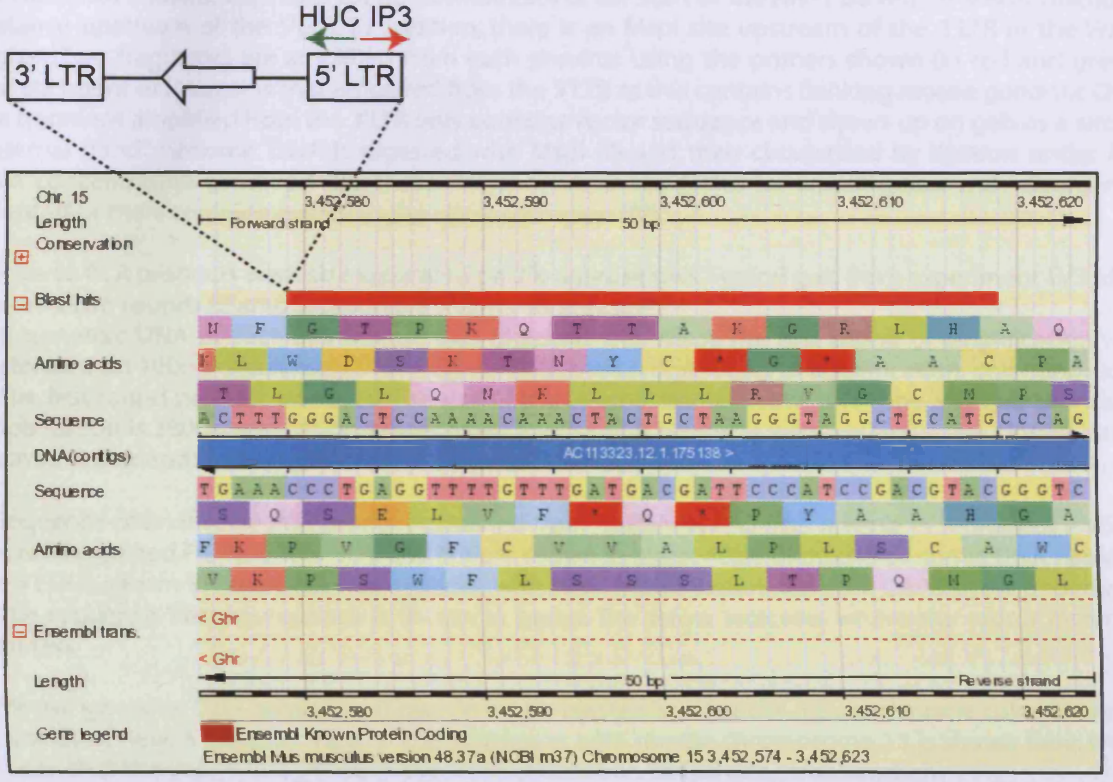
IP3

← TCACTCCCAACGAAGACAAGA

CCCTGGAAGGGCTAATTCACTCCCAACGAAGACAAGA

Junction

d.



Query location	:	unnamed	1 to 40 (-)
Database location	:	AC113323.12.1.175138	22062 to 22101 (+)
Genomic location	:	15	3452579 to 3452618 (+)

Alignment score	:	40
E-value	:	8.0e-15
Alignment length	:	40
Percentage identity:		100.00

Figure 4.3. c. and d.

Figure 4.3. HV lentiviral vector integration sites in IL-3 independent mutants were cloned using the inverse PCR technique

The inverse PCR technique is applied to amplify DNA of unknown sequence that is adjacent to DNA of known sequence. Silver et al first described the application of this technique for the amplification of cellular DNA adjacent to integrated provirus. The inverse PCR protocol shown here was adapted from that described by Carteau et al.

a. Overview of the inverse PCR technique. The HV lentiviral vector is shown as a provirus, flanked by mouse genomic DNA of unknown sequence in blue (I). Genomic DNA was digested by 4bp-cutter MspI (or NlaIII, not shown). MspI cuts 237bp downstream of the start of the HIV-1 U3 region and at unknown distance upstream of the 5'LTR. In addition, there is an MspI site upstream of the 3'LTR in the WPRE region. Two fragments are amplified from each provirus using the primers shown (in red and green). The fragment of interest is that amplified from the 5'LTR as this contains flanking mouse genomic DNA; the fragment amplified from the 3'LTR only contains vector sequence and shows up on gels as a strong "internal band". Genomic DNA is digested with MspI (II) and then circularised by ligation under low DNA concentration condition (III). These circles serve as templates for amplification with divergently oriented primers complementary to the HIV-1 U3 region (IV).

b. Inverse PCR products were size separated on 2% agarose gels. Typical gels from experiment LV3 after one and two rounds of amplification are shown, as indicated.
2μg genomic DNA of each LV3 mutant was digested with MspI. The first round of amplification was performed on 100ng circularised DNA fragments. In the second round of amplification, one hundredth of the first round product was used. The size of the internal band (arrow) after one and two rounds of amplification is 380bp and 345bp, respectively. Inverse PCR products were gel extracted, cloned into a plasmid vector and then sequenced.

c. Sequence of an inverse PCR product obtained from mutant HV A2. The inverse PCR product is 95bp in size. The nested PCR primers HUC and IP3 are shown in black. 40bp of mouse genomic DNA flanking the 5'LTR is shown in blue. DNA was digested with NlaIII in this experiment; the CATG site is underlined in the sequence. Vector sequence is shown in green. The arrow indicates where the vector insertion occurred.

d. Mouse genomic DNA sequence shown in c. was blasted against the mouse genome using Ensembl Mouse BlastView. A basepair view of the alignment with mouse chromosome 15 is shown here. It can be seen that the vector integrated in the same orientation as Ghr gene transcription.

4.2.4.2. Identification of HV lentiviral vector integration sites in IL-3 independent mutants

To investigate why the mutants became IL-3 independent, the genomic locus/loci where the vector(s) had integrated needed to be identified. The Southern blots shown in Figure 4.2. gave a good estimate of how many vector integration sites would need to be cloned out in each of the mutants. Mouse genomic DNA flanking the 5'LTRs of HV vector proviruses in the mutants from LV1-3 was amplified by inverse PCR. The principle of this method is shown in Figure 4.3.a. Inverse PCR allows amplification of DNA of unknown sequence that is adjacent to DNA of known sequence. Mutant genomic DNA was digested with a 4bp-cutter restriction enzyme. Digested DNA is then ligated under low DNA concentration condition that favours intramolecular ligation i.e. circularisation. This is used as template in PCR reactions using divergently oriented primers, complementary to the HIV-1 U3 region. In mutants that contained several copies of the vector, inverse PCR was performed using at least two different restriction enzymes, so most vector integration sites could be obtained. In most experiments I was able to clone out about half the total number of vector integration sites by inverse PCR. Inverse PCR products were size separated on agarose gels (Figure 4.3.b), extracted, cloned into a plasmid vector and then sequenced. The integrity of any sequences was assessed by checking for the presence of both primers, the restriction site, the 5bp duplication upstream of U3 and the appropriate length of vector sequence (75bp when MspI digested; 60bp when NlaIII digested)(Figure 4.3.c). Sequences determined to be genuine integration sites were then blasted against the mouse genome (Figure 4.3.d). The 25 HV vector integration sites recovered in mutants from LV1-3 are shown in Table 4.2. No integration sites were recovered from mutants HV10, HV25, HV26 and HV42. Strikingly, in 10 IL-3 independent mutants integration sites into the Growth Hormone Receptor (Ghr) locus on mouse chromosome 15 were found. This locus was previously identified as a retroviral common insertion site, Evi124. It is listed 6 times in the Retrovirus Tagged Cancer Gene Database (Akagi *et al.* 2004). Mutants HV A2 and HV49 each contain just a single copy of the vector. In both these mutants, the vector

had integrated into the Ghr, suggesting activation of this gene alone is sufficient to transform BAF3/BCL15 cells. Integration into the Plekha5 gene on mouse chromosome 6 was also found in two different mutants, HV3 and HV48.

4.2.4.3. Locations of HV vector insertions into the Growth Hormone Receptor locus

The locations of the HV vector insertions into the Ghr locus in the 10 different IL-3 independent mutants are shown in Figure 4.4. All HV vector insertions occurred in intron 1-2 of the Ghr, in the same transcriptional orientation as Ghr transcription. Intron 1-2 of the Ghr is over 125 kilobases in length; the first coding exon is exon 2. The mouse Ghr uses at least three alternative promoters and associated first exons (Menon *et al.* 1995; Yu *et al.* 1999; Menon *et al.* 2001). The transcripts originating from these promoters differ only in their 5' untranslated regions and are termed L1, L2 and L5. Ghr promoter usage is tissue-specific and also depends on the developmental stage of the organism. Most Ghr transcripts in all tissues originate from the L2 promoter, which is 125 kilobases upstream of exon 2 (Edens and Talamantes 1998). The L1 promoter is predominantly used in pregnant liver (Schwartzbauer *et al.* 1998). The L5 promoter and first exon are immediately upstream of exon 2. The L5 transcript is most abundant in foetal liver and placenta where it makes up 10-15% of Ghr transcripts (Menon *et al.* 2001). As can be seen in Figure 4.4., most HV vector insertions occur in a 15 kilobase region immediately upstream of exon 2. They are close to the L1 or L5 promoters. Only the vector integration site in Mutant HV A2 is not near any of the three known alternative promoters.

Table 4.2. HV lentiviral vector integration sites in IL-3 independent mutants (LV1-3)

Clone	SB	IPCR	In Gene (Y/N)	Gene Name	Entrez Gene ID	Chr	Distance to TSS	Ori	Predicted Gene Function	RTCGD
LV1										
HVA2	1	1	Y	Ghr	14600	15		+	*Growth Hormone Receptor*	6 Evi124
LV2										
HV3	12	6	Y	Ghr	14600	15		+	*Growth Hormone Receptor*	6 Evi124
			Y	Dock2	94176	11	323kb Intron 27-28	+	Cytoskeletal organisation	
			Y	Stag1	20842	9	149kb Intron 7-8	+	Cell cycle; cell division	2 Evi99
			Y	Hcfc2	67933	10	5867bp Intron 3-4	-		
			Y	Plekha5	109135	6	103kb Intron 5-6	-		
			Y	2610020 H08 Rik	434234	7	48kb Intron 12-13	+		
HV14	7	2	Y	Ghr	14600	15		+	*Growth Hormone Receptor*	6 Evi124
			Y	Usp45	77593	4	34kb intron 9-10	+	Ubiquitin cycle	
LV3										
HV24	3	2	Y	Ghr *	14600	15		+	*Growth Hormone Receptor*	6 Evi124
			N			14				
HV32		2	Y	Ghr	14600	15		+	*Growth Hormone Receptor*	6 Evi124
			N			7	37kb down-stream of Btbd1			
HV41	4	2	Y	Ghr	14600	15		+	*Growth Hormone Receptor*	6 Evi124
			Y	Suclg2	20917	6	26kb Intron 1-2	+	Metabolic process; GDP forming activity; tricarboxylic acid cycle	
HV43	4	1	Y	Ghr	14600	15		+	*Growth Hormone Receptor*	6 Evi124

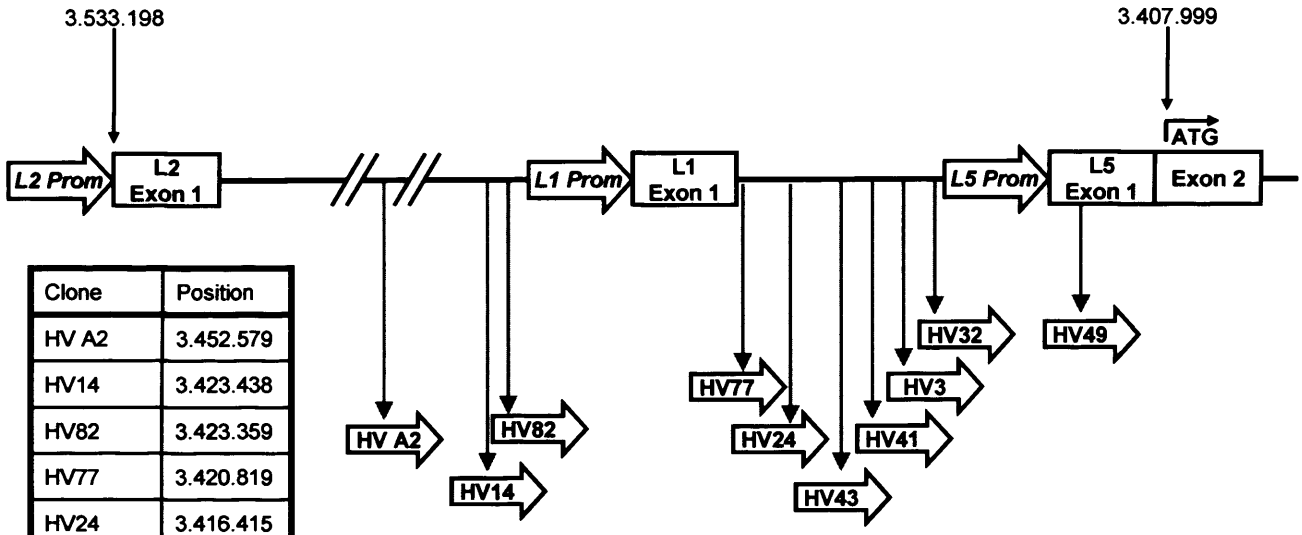
Clone	SB	IPCR	In Gene (Y/N)	Gene Name	Entrez Gene ID	Chr	Distance to TSS	Ori	Predicted Gene Function	RTCGD
HV48		2	Y	Senp1	223870	15	7kb Intron 2-3	-	Multicellular organismal dev't; protein desumoylation	
			Y	Plekha5	109135	6	3kb Intron 1-2	+		
HV49	1	1	Y	Ghr	14600	15		+	*Growth Hormone Receptor*	6 Evi124
HV72	2	1	Y	Ccdc111	408022	8	10kb Intron 4-5	-	DNA replication	
HV77	4	4	Y	Ghr *	14600	15			*Growth Hormone Receptor*	6 Evi124
			Y	Gmps	229363	3	1217bp Intron 1-2	+	Purine nucleotide biosynthetic process; glutamine metabolic process	
			N			16	39kb upstream of Alcam			
			Y	Hif1a	15251	12	15kb Intron 8-9	+	Angiogenesis; cartilage dev't	
HV82		1	Y	Ghr	14600	15		+	*Growth Hormone Receptor*	6 Evi124

Table 4.2. HV lentiviral vector integration sites in IL-3 independent mutants (LV1-3)

Inverse PCR was performed on IL-3 independent mutants from LV1-3 using two different restriction enzymes, MspI and NlaII. PCR products were gel extacted, cloned into a plasmid vector and then sequenced. Vector-genome junction sequences were blasted against the mouse genome, released August 2007 (www.ensembl.org/Mus_musculus).

Some vector integration sites were obtained in the majority of mutants, with the exception of LV3 mutants HV10, HV25, HV26 and HV42.

SB indicates vector copy number in the mutant determined by Southern blot. IPCR indicates the number of insertion site loci identified in the mutant by inverse PCR. The Ghr insertion sites in mutants HV24 and HV77 (asterisks) were cloned out by multiplex PCR. All Ghr insertion sites were confirmed by site-specific PCR. For most insertion sites the distance to the transcription start site (TSS) of the nearest gene is shown, except in the case of Ghr insertions because this gene uses several TSS. The locations of HV vector insertions into the Ghr are shown in Figure 4.4. Ori indicates whether the vector insertion occurred in the same (+) or in the opposite (-) orientation as the gene's transcript. Predicted gene function is shown where known. The number of appearances in the retroviral tagged cancer gene database (RTCGD) of each insertion site locus is shown where applicable.



Clone	Position
HV A2	3.452.579
HV14	3.423.438
HV82	3.423.359
HV77	3.420.819
HV24	3.416.415
HV43	3.414.174
HV41	3.411.498
HV3	3.409.827
HV32	3.409.089
HV49	3.408.558

Figure 4.4. Location of HV vector insertions into the growth hormone receptor (Ghr) locus in 10 IL-3 independent mutants.

All 10 HV vector insertions occurred in the same orientation as Ghr gene transcription and were in the 125 kilobasepair region upstream of exon 2, the first coding exon of the Ghr. The three alternative promoters and associated first exons used by the Ghr, are also shown. ATG indicates the start codon in exon 2. The table shows the position on mouse chromosome 15 where each vector integration occurred.

4.2.4.4. All Ghr insertional mutants express the Ghr transcript

To investigate whether the Ghr transcript was expressed in the IL-3 independent mutants that had HV vector insertions into the Ghr locus, reverse transcriptase (RT) PCR for the Ghr transcript was performed (Figure 4.5.a.). The Ghr transcript was indeed expressed in all 10 mutants; in addition it was also produced by mutants HV10, HV25, HV26, HV42 and HV43. Not all HV vector integration sites in these mutants were cloned out by inverse PCR; presumably they also have vector insertions into the Ghr. Insertional activation of the Ghr therefore seems to be the most common mechanism whereby the HV lentiviral vector transforms BAF3/BCL15 cells to IL-3 independence. The Ghr transcript is not expressed by the parental cells, mock mutants M8 and M53 and mutant HV48. However, mutant HV48 does express the IL-3 transcript, as do the two mock mutants (Figure 4.5.b). Activation of IL-3 gene expression seems to be a mechanism whereby BCL15 cells spontaneously become IL-3 independent.

4.2.4.5. Ghr-insertional mutants express the Growth Hormone Receptor on the cell surface

Insertional activation of the Ghr leads to expression of a functional GHR on the cell surface of IL-3 independent mutants. Cell surface staining for the GHR was performed on mutants HV3 and HV14 from experiment LV2. Unlike the parental cells, the mutants express GHR on the cell surface (Figure 4.6.). Dr E.F. Gevers at the National Institute for Medical Research in Mill Hill demonstrated that the GHR on these mutants was functional by performing intracellular staining for phosphorylated STAT5 (Data not shown). STAT5 was phosphorylated in parental cells and mutants following stimulation with IL-3; bGH stimulated STAT5 phosphorylation in mutants only.

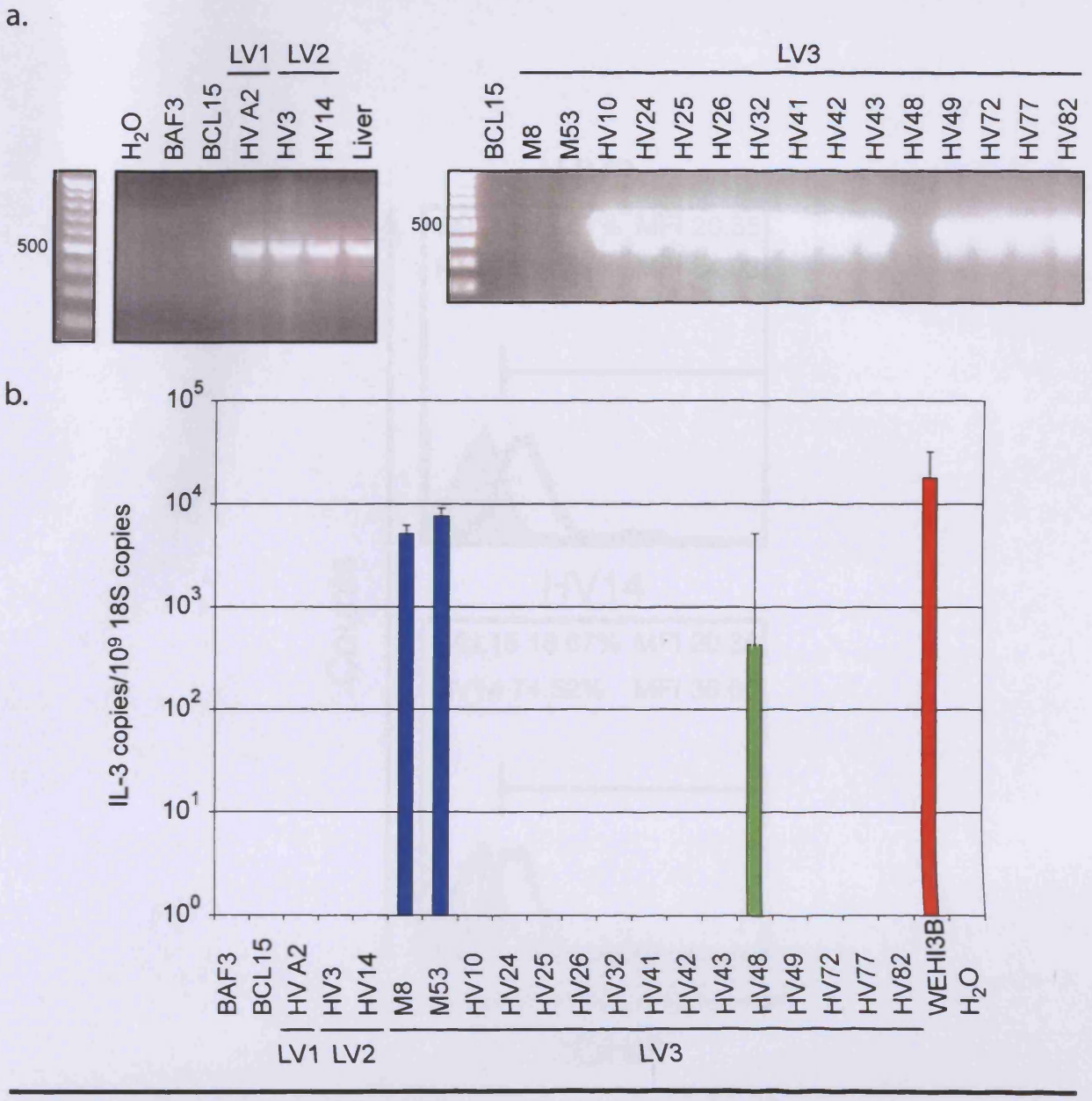


Figure 4.5. All 10 mutants that contain an HV vector insertion into the Ghr locus, express the Ghr transcript.

a. A reverse-transcriptase (RT) PCR for expression of the Ghr transcript was performed on parental cells and IL-3 independent mutants from experiments LV1-3. A forward primer in exon 4 and a reverse primer in exon 8 were used to amplify a 552bp section of the GHR transcript. Mouse liver RNA was used as a positive control. The 10 mutants that were found to have HV vector insertions into the Ghr, all express the Ghr transcript. Mutants HV10, HV25, HV26, HV42 and HV72 also produce the Ghr transcript. These mutants presumably have vector integrations into the Ghr, though none were recovered by inverse PCR. Parental BAF3 and BCL15, the two mock mutants and mutant HV48 do not express the Ghr transcript.

b. Quantitative RT-PCR to measure IL-3 transcript levels was performed. Results are expressed as IL-3 copies per 10^9 18S copies. The error bars represent the standard error of this ratio. As positive control, RNA from WEHI3B cells was used. The two mock mutants and mutant HV48 express the IL-3 transcript. This suggests that autocrine IL-3 production is a cause of spontaneous IL-3 independence in BCL15 cells.

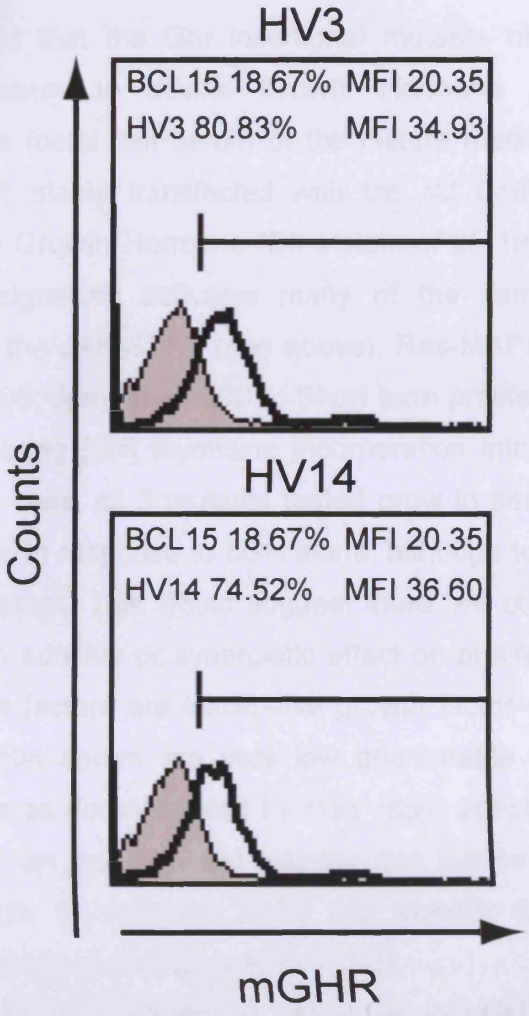


Figure 4.6. Mutants HV3 and HV14 from experiment LV2 express the Growth Hormone Receptor (GHR) on the cell surface.

Using an antibody that recognises the extracellular domain of the mouse GHR, it was demonstrated that unlike parental BCL15 cells (solid gray peaks) mutants HV3 and HV14 (black peaks) express the mouse GHR on the cell surface. The percentage of cells to the right of the gate and the mean fluorescence intensity (MFI) of the peaks are shown in each histogram.

4.2.4.6. Ghr-insertional mutants respond to bovine Growth Hormone in culture

It was hypothesised that the Ghr insertional mutants obtained in the assay, proliferate in response to bovine Growth Hormone present at unknown concentration in the foetal calf serum of the culture medium. It was previously reported that BAF3 stably transfected with the rat GHR could proliferate in response to bovine Growth Hormone (Dinerstein *et al.* 1995; Jeay *et al.* 2000). Growth hormone signalling activates many of the same pathways as IL-3 signalling, such as the JAK-STAT (see above), Ras-MAPK and PI3K pathways (Gouilleux *et al.* 1995; Jeay *et al.* 2001). Short term proliferation of mutants was assessed by measuring [3H] thymidine incorporation into DNA (Figure 4.7.a.). Unlike the parental cells, all 3 mutants tested grow in serum only without IL-3. They also proliferate in response to bGH alone, although to a lesser degree than when cultured in serum. This would suggest there are other factors present in serum that have an additive or synergistic effect on proliferation in combination with GH. Candidate factors are insulin-like growth factor 1 (IGF-1) and insulin. IGF-1 levels in foetal serum are very low and unable to support growth of parental BAF3 cells as demonstrated by their rapid apoptosis when cultured in serum only. IGF-1 can maintain cell viability and induce proliferation of BAF3 cells; insulin is able to maintain BAF3 cell viability but only at very high concentrations (Rodriguez-Tarduchy *et al.* 1992).

Ghr-insertional mutants proliferate in response to bGH in a dose-dependent manner; proliferation is maximal at bGH concentrations of 1 µg/ml or above (Figure 4.7.b.). Ghr-insertional mutants are unable to proliferate in the complete absence of serum. Long term proliferation of mutants in response to bGH was also assessed by counting trypan blue excluding cells in cultures for the duration of one week (Figure 4.7.c). Interestingly, all 3 mutants proliferate faster in serum without IL-3 than in serum with IL-3; the difference in population doubling time is approximately two-fold.

Both parental cells and mutants rapidly lose viability when cultured in serum-free medium. In contrast to the data shown in Figure 4.7.a, I could not observe an increase in cell number if HV A2 mutants were cultured in bGH only. Cells remained viable for 72 hours, followed by a decrease in cell number. This was also seen when HV A2 were cultured with 100 ng/ml IL-3, suggesting that neither signal maintains Bcl-X_L expression effectively. HV3 cells cultured in the presence of bGH or IL-3 only, proliferated for up to 4 days at a similar rate to HV3 cells grown in serum plus WEHI. The total cell number slowly dropped thereafter. This was also seen for HV14 cells cultured with bGH or IL-3 only, though in these cultures the cell number dropped after 72 hours. Therefore, neither GH nor IL-3 alone is able to sustain long term proliferation of BAF3 or BCL15 cells at these concentrations.

4.2.4.7. Proliferation of GHR-expressing mutants seems not to be mediated by IGF-1 induction

Conflicting reports exist in the literature with regards to IGF-1 production in response to GH signalling by GHR-expressing BAF3 cells. Proliferation of BAF3 cells stably expressing the rat GHR was shown to be a direct effect of GH-signalling. Following GH stimulation of these cells, IGF-1 mRNA could not be detected by RT-PCR, neither could IGF-1 protein be detected in the culture medium (Baixeras *et al.* 2001). However, another study that analysed human GHR expressing BAF3 cells showed IGF-1 expression in these cells was stimulated by human GH in a dose-dependent fashion (Yoshizato *et al.* 2000). I could not detect IGF-1 gene expression by RT-PCR in my GHR-expressing insertional mutants that were grown in serum (Data not shown).

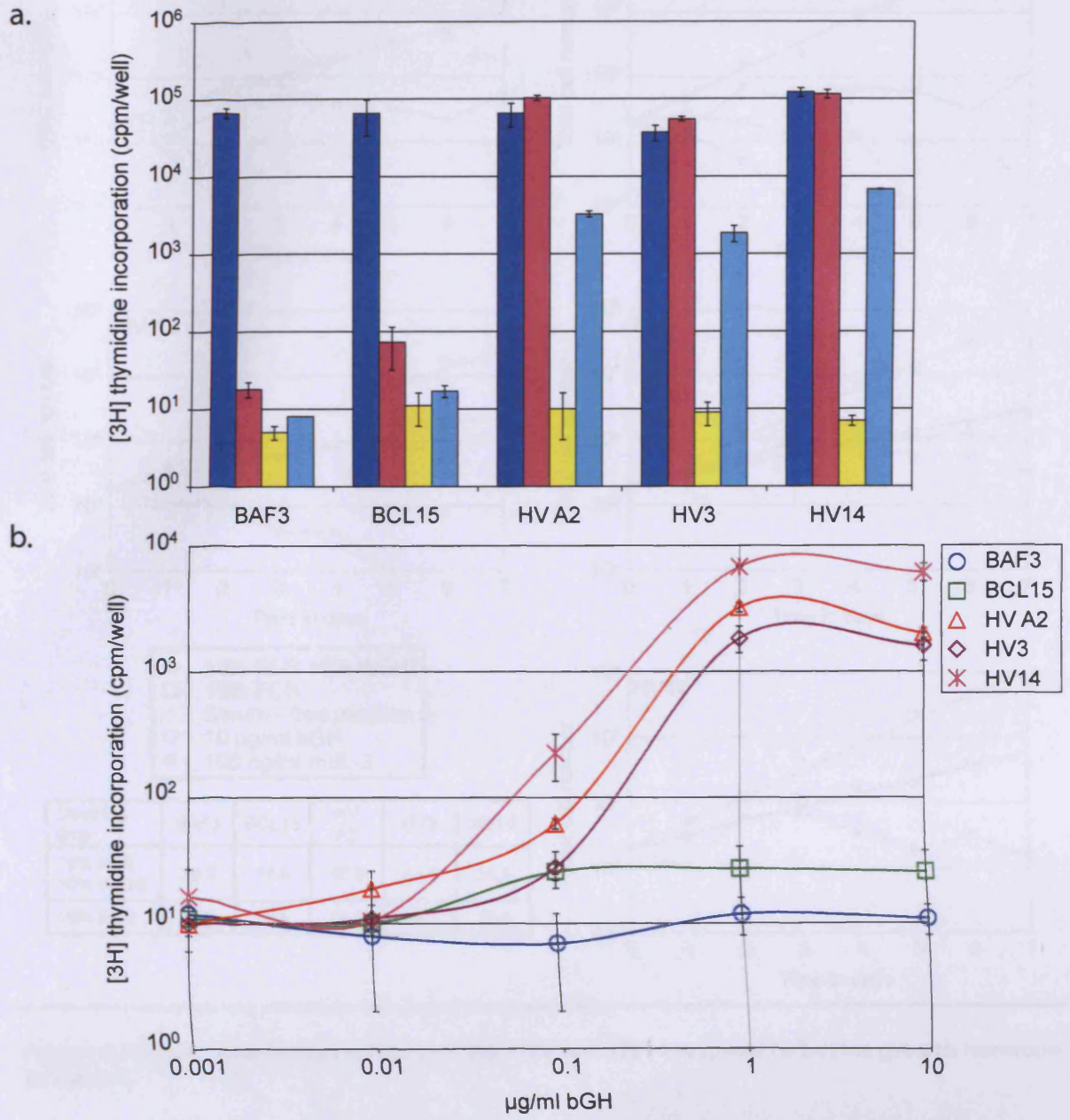
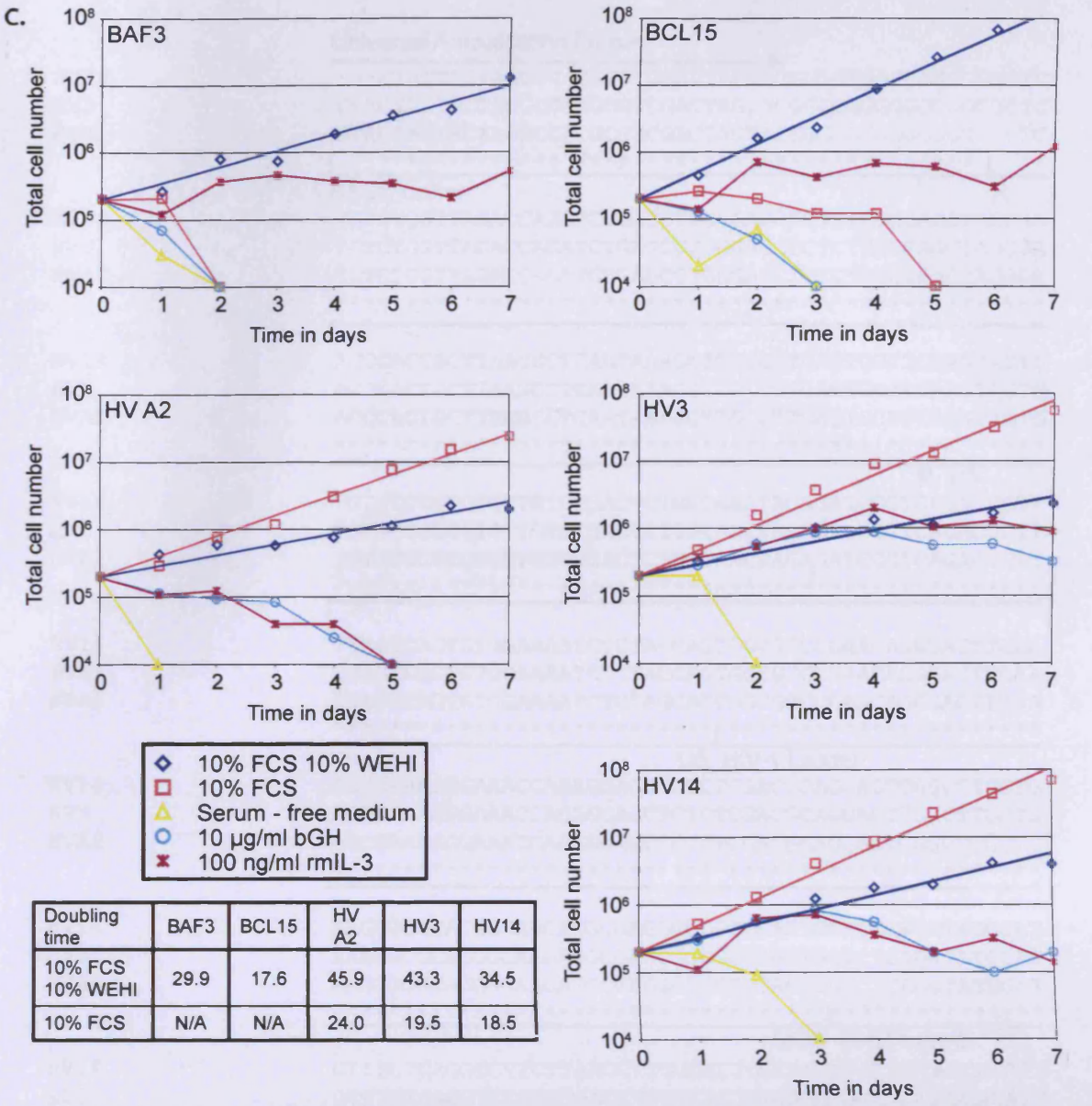


Figure 4.7. Ghr insertional mutants HV A2, HV3 and HV14 respond to bovine growth hormone in culture.

a. Parental cells and mutants were cultured in: medium containing 10% FCS and 10% WEHI (blue); medium containing 10% FCS only (pink); serum-free medium (yellow) or serum-free medium containing 1 µg/ml recombinant bGH (aqua blue).
Cells were cultured in these media for 24 hours before an overnight incubation in the presence of [3H] thymidine. [3H] thymidine incorporation was measured by liquid scintillation counting the following day. Error bars indicate the standard error of the mean calculated from 2 independent experiments.
b. Parental cells and mutants were incubated with serum-free medium containing different concentrations of bGH; DNA synthesis was measured as above.

Ghr insertional mutants proliferate in response to bGH (a). This response was shown to be bGH dose-dependent (b). It plateaus at bGH concentrations from 1 µg/ml. Mutant proliferation in response to bGH alone is however less than in the presence of FCS (a), suggesting there are other mitogens present in serum that cause these cells to grow.



Doubling time	BAF3	BCL15	HV A2	HV3	HV14
10% FCS 10% WEHI	29.9	17.6	45.9	43.3	34.5
10% FCS	N/A	N/A	24.0	19.5	18.5

Figure 4.7.c. Ghr insertional mutants HV A2, HV3 and HV14 respond to bovine growth hormone in culture.

Long term proliferation of Ghr-insertional mutants in response to bGH was assessed by counting trypan blue excluding cells every day for one week.
2 x 10^5 cells were cultured in 2ml volume in: medium containing 10% FCS and 10% WEHI (blue); medium containing 10% FCS only (pink); serum-free medium (yellow); serum-free medium containing 10 µg/ml recombinant bGH (aqua blue) or serum-free medium containing 100 ng/ml rmIL-3 (purple).
When the cell count approached 10^6 cells/ml, cells were passaged. Every two days, at 48, 96 and 144 hours, 1ml medium was replaced by fresh medium in all wells.
In the charts, the total cell number is plotted against time, taking cell passages into account. Exponential trendlines were fitted to all "10% FCS plus 10% WEHI" data series. R-squared values were 0.969, 0.9931, 0.9027, 0.8789 and 0.9549 for BAF3, BCL15, HV A2, HV3 and HV14, respectively. Exponential trendlines were also fitted to the "10% FCS" data series of the 3 mutants; R-squared values were 0.9882, 0.9814 and 0.9875 for HV A2, HV3 and HV14, respectively. The population doubling time, shown in the table, was calculated using the formula $y = c * e^{rx}$, where c is the intercept (set at 2 x 10^5), e is the base of the natural logarithm and r is the growth constant. For the "10% FCS plus 10% WEHI" data series the growth constants were 0.0232, 0.0394, 0.0151, 0.016 and 0.0201 for BAF3, BCL15, HV A2, HV3 and HV14, respectively. For the "10% FCS" data series the growth constants were 0.0289, 0.0356 and 0.0375 for HV A2, HV3 and HV14, respectively.

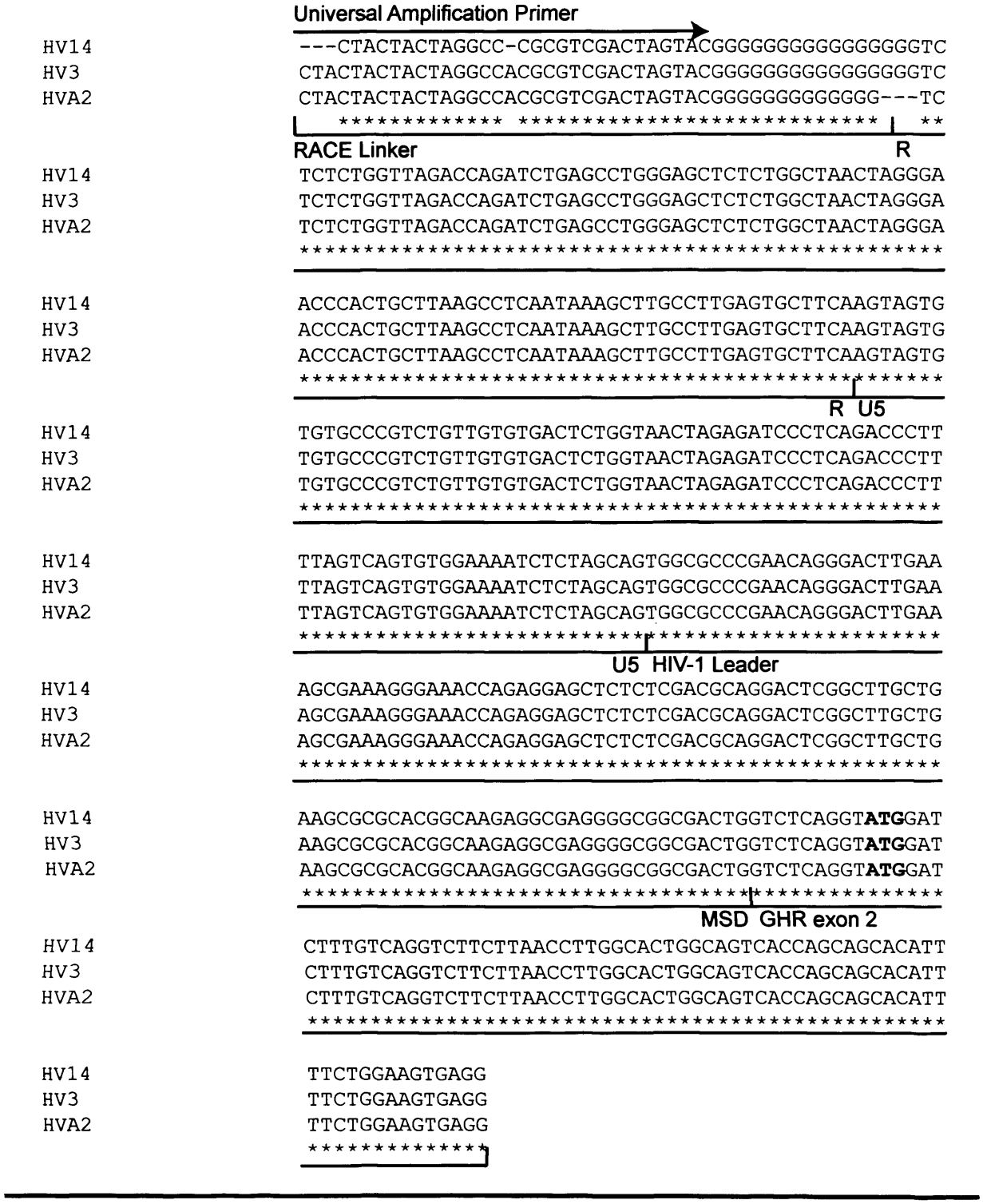


Figure 4.8. The Ghr transcript in the mutants originates from the HIV-1 5'LTR of the vector.

To amplify the 5' end of the Ghr transcript present in the mutants, Dr S.L. Stephen in the Collins lab performed 5' rapid amplification of cDNA ends (5' RACE) on mutants HV A2, HV3 and HV14. cDNA was synthesised from total RNA using a Ghr-specific reverse primer that anneals in exon 5. An oligo-dC tail was added to the 3' end of this cDNA using the enzyme terminal deoxynucleotidyl transferase (TdT). Nested amplification of tailed cDNA was performed using universal forward primers whose 3' sequence is complementary to the oligo-dC tail and Ghr-specific reverse primers in exon 4. A 590bp 5' RACE product was cloned and sequenced. An alignment of the first 414bp of the 5'RACE product is shown.
The RACE linker cassette is 46bp; the GHR transcript starts at the HIV-1 R region. It contains 289bp of HV vector sequence, before it splices from the HIV-1 major splice donor to the splice acceptor of GHR exon 2. The start codon in Ghr exon 2 is highlighted in bold.

4.2.4.8. Analysis of the Ghr transcript in IL-3 independent mutants

To identify the mechanism whereby HV vector insertion activates Ghr gene expression, we needed to analyse the Ghr transcript in the mutants. This work was carried out by Dr S.L.Stephen in the Collins lab. To investigate where the Ghr transcript originates from, 5'RACE was performed. This specialised form of RT-PCR allows amplification of the 5'ends of rare mRNA species. 5'RACE on mutants HV A2, HV3 and HV14 yielded a single pre-dominant PCR product of approximately 600bp in size. This 5'RACE product was cloned and sequenced. An alignment of the sequences is shown in Figure 4.8. The Ghr transcript starts at the beginning of the HIV-1 R region; it further contains the 5'LTR U5 region and HIV-1 leader sequence up to the HIV-1 major splice donor (MSD); the Ghr transcript splices from the HIV-1 MSD to the splice acceptor of Ghr exon 2. Further RT-PCR analysis by Dr S.L.Stephen failed to detect Ghr transcripts that contained any of the three known alternative first exons (L2/L1 and L5) or the HIV-1 U3 region (Data not shown). Hence, the predominant Ghr transcript in Ghr-insertional mutants is a vector-Ghr fusion transcript that seems to originate from the HIV-1 5'LTR.

4.2.5.The frequency at which Ghr insertional mutants are obtained from an HV vector transduced population can be increased by selecting cells in serum plus additional growth hormone

So far, Ghr insertional mutants were only obtained following selection according to the standard selection protocol, whereby cells were grown in 24 well plates in complete medium without IL-3 for one week; then rescued by adding IL-3 back to the medium; then washed of IL-3 again for a second selection round for IL-3 independence. Ghr insertional mutant clones were unable to grow out without IL-3 rescue when selected in 24 well plates in complete medium, even though established mutants proliferate very well in this medium. We were wondering whether cells in which the HV vector had activated the Ghr could expand when selected in complete medium containing additional bGH. A modified protocol whereby cells are selected in 24 well plates in complete medium containing 1

μg/ml bGH was first tested in experiment LV3. Its efficiency in selecting for Ghr insertional mutants was compared to that of the standard selection protocol. Equal numbers of HV transduced target cells in experiment LV3 were selected according to each method. 26 out of 100 wells containing growing cells were scored after selection in serum plus bGH. After elimination of "sibling" clones, another 18 independent mutants with the HV vector were obtained. This was at a cell and integrant frequency of 3.6×10^{-7} and 9.5×10^{-8}, respectively (Table 4.3.). No mock mutants were obtained in this experiment. All these new "GH HV" mutants were able to proliferate in complete medium only without the need for additional bGH once they had become established clones. We wanted to determine whether any of the same Ghr insertional mutants were picked up after selection with additional GH compared to standard selection. Southern blots on the 18 "GH HV" mutants were performed (Figure 4.9.) and compared to the Southern blots shown in Figure 4.2. of the LV3 standard selection mutants. No mutants which shared band patterns could be seen. It therefore looks like 31 unique mutants were obtained from the initial HV-transduced population in experiment LV3. HV vector integration sites in the "GH HV" mutants were not identified, but all expressed the Ghr transcript (Figure 4.10.a) and none expressed the IL-3 transcript (Figure 4.10.b). Selecting HV vector transduced cells in complete medium with additional GH is therefore an effective way to select for Ghr insertional mutants.

Table 4.3. HV lentiviral vector mutagenesis frequencies after standard selection and selection in foetal calf serum plus 1 µg/ml bovine growth hormone (LV3).

Selection Type	Target Cell	Mock Mutants		HV Vector Transduced Mutants		
		Mutant Number	Cell Frequency [a]	Mutant Number	Cell Frequency [a]	Integrant Frequency [b]
Standard	BCL15	2	4.0×10^{-8}	13	2.6×10^{-7}	1.8×10^{-7} 6.8×10^{-8}
FCS + bGH	BCL15	0	$< 2.0 \times 10^{-8}$	18	3.6×10^{-7}	2.5×10^{-7} 9.5×10^{-8}

Table 4.3. Experiment LV3. HV lentiviral vector mutagenesis frequencies after standard selection and selection in foetal calf serum plus 1 µg/ml bovine growth hormone.

In experiment LV3, 5 x 10^7 BCL15 target cells were transduced with HV lentiviral vector at MOI 3.77 (Q-PCR). Cells were cultured in the presence of IL-3 for 48 hours following transduction. By day 3 of the experiment, target cells had expanded 8-10 fold in cell number. 10^8 cells were washed thrice in complete medium without IL-3. 5 x 10^7 cells each were selected in one of two ways:

a. Selection according to the standard protocol as described in Table 4.1. Briefly, cells are plated in 24 well plates at 5 x 10^5 cells per well in 2ml complete medium without IL-3. Cells are rescued by adding IL-3 back to the medium on Day 10. Any growing cells are washed off IL-3 again to subject them to a second selection round for IL-3 independence.

b. Selection in complete medium supplemented with 1 µg/ml bGH in 24 well plates at 5 x 10^5 cells per well in 2ml medium. Wells are fed once a week by replacing 1ml old medium for 1ml fresh medium containing bGH. Plates were screened by visual inspection for cells growing out.

[a] The cell frequency was calculated by dividing the number of mutants by the total number of target cells in the experiment. The difference between the observed cell frequencies in the mock and HV arms of the experiment was significant: LV3 standard selection $p < 0.001$, $X^2 = 161$; LV3 FCS + bGH selection $p < 0.001$; $X^2 = 360$.

[b] The integrant frequency was calculated by dividing the number of mutants by the total number of integrants in each experiment.

The number of "GFP" integrants was calculated by multiplying the number of live target cells by the MOI. The MOI was determined by flow cytometry for GFP expression at 48h to 72h post-transduction. The GFP integrant frequency is shown in green.

The average number of vector copies per cell in experiment LV3 was also determined by Q-PCR. This number was multiplied by the number of target cells to obtain the number of "Q-PCR" integrants. The Q-PCR integrant frequency is shown in black.

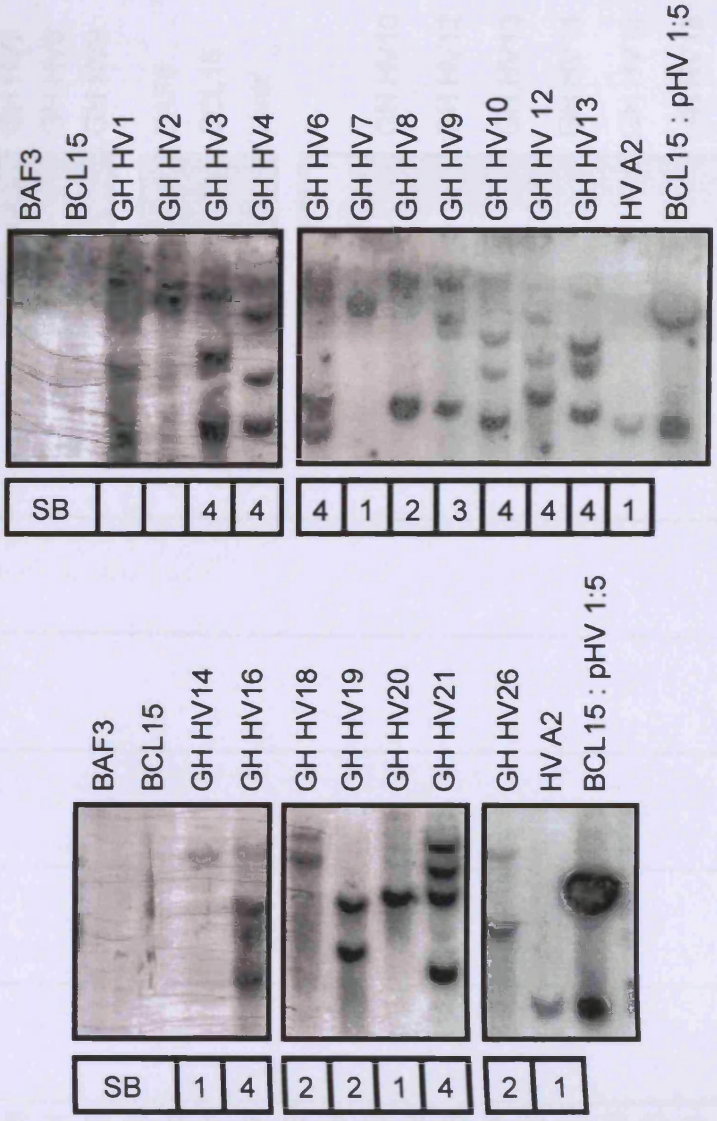


Figure 4.9. Southern blots of IL-3 independent mutants obtained in experiment LV3 following selection in foetal calf serum supplemented with 1µg/ml bGH.

Southern blots were performed as described in Figure 4.2. Band patterns were examined to determine clonality of the mutants and HV vector copy number per genome. Blots from LV3 were exposed to autoradiography film. Irrelevant lanes on blots were removed. The number of bands seen on the Southern blots of each mutant is shown in the table below the blot.
Southern blot band patterns of LV3 mutants selected in complete medium plus bGH (shown in this figure), were compared to those of LV3 mutants obtained by standard selection (shown in Figure 4.2.). No band pattern similarities could be seen, suggesting unique mutants were picked up by both selection methods.

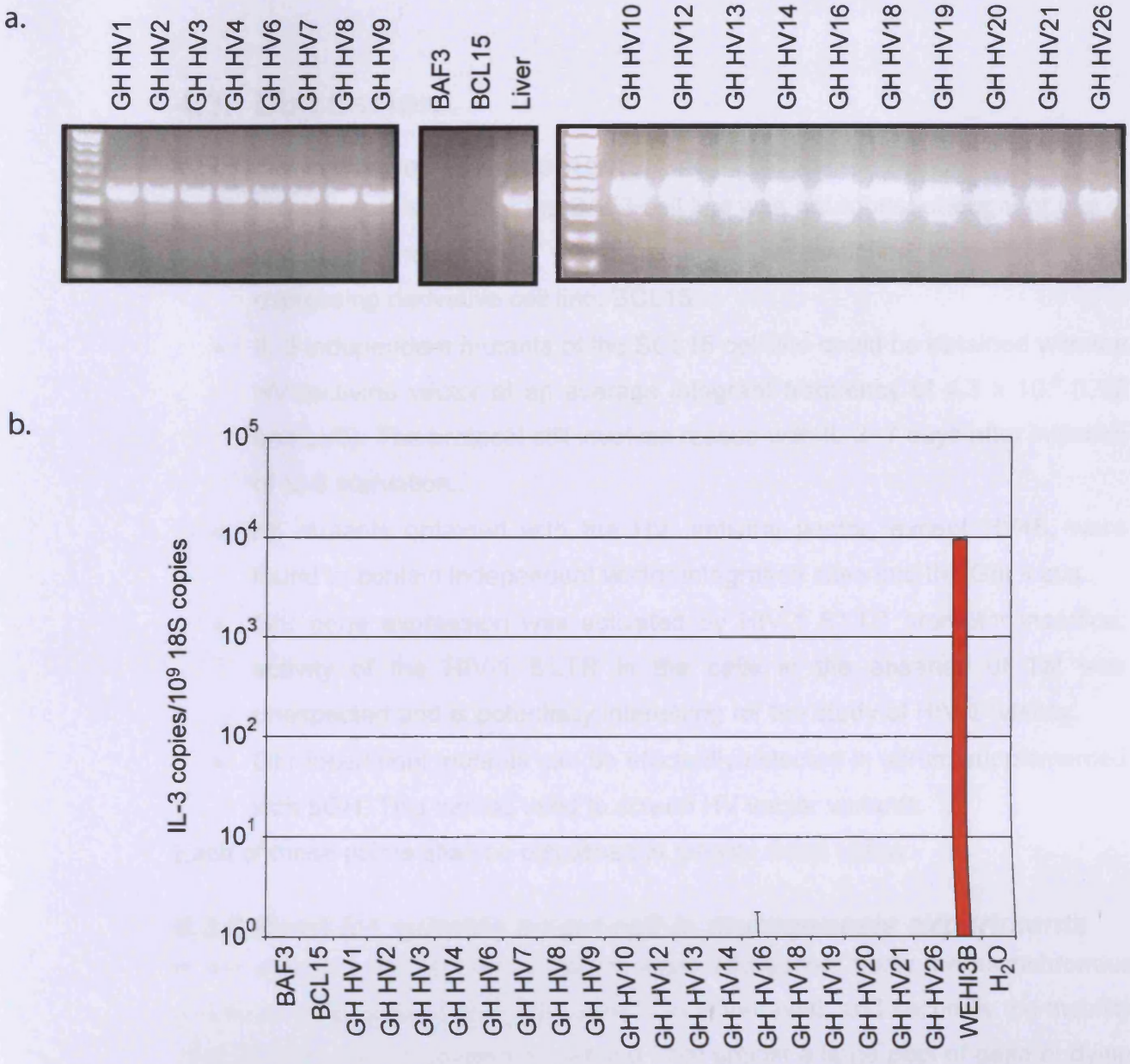


Figure 4.10. IL-3 independent mutants that express the Ghr transcript, are obtained following selection of an HV vector transduced population in serum plus additional bovine GH.

RT-PCR for expression of the Ghr transcript (a) and quantitative RT-PCR to measure IL-3 transcript levels (b) was performed as described in Figure 4.5.
IL-3 independent mutants from experiment LV3 that were selected in serum plus additional bGH all express the Ghr transcript (a); none express the IL-3 transcript (b).

4.3. Discussion

4.3.1. Summary of achievements

- It was established that the BAF3 cell line was not robust enough for use in insertional mutagenesis experiments, so we switched to a Bcl-2 over-expressing derivative cell line, BCL15.
- IL-3 independent mutants of the BCL15 cell line could be obtained with the HV lentiviral vector at an average integrant frequency of 4.3×10^{-8} (LV2 and LV3). The protocol still involves rescue with IL-3, 7 days after initiation of IL-3 starvation.
- All mutants obtained with the HV lentiviral vector, except HV48, were found to contain independent vector integration sites into the Ghr locus.
- Ghr gene expression was activated by HIV-1 5'LTR promoter insertion; activity of the HIV-1 5'LTR in the cells in the absence of Tat was unexpected and is potentially interesting for the study of HIV-1 latency.
- Ghr-insertional mutants can be efficiently selected in serum supplemented with bGH. This can be used to screen HV vector variants.

Each of these points shall be discussed in greater detail below.

4.3.2. Need for suitable target cell in mutagenesis experiments

In the previous chapter, two problems were addressed: firstly the asynchronous cell death of parental BAF3 cells upon IL-3 withdrawal, and secondly the inability of IL-3 independent mutants to expand from amidst a large pool of dead or dying cells, necessitating rescue with IL-3. This chapter started by outlining a third problem, namely the inability to generate IL-3 independent mutants of the BAF3 cell line with a lentiviral vector that had been shown in a pilot experiment to be able to transform this cell line. Several variables were tested: selection in T25 flasks versus 24 well plates; cell density during selection; whether or not to rescue by adding IL-3 back and if IL-3 rescue was performed, what is the best time to do this following initiation of IL-3 starvation. However, no IL-3 independent mutants of the BAF3 cell line could be obtained.

At the time this left us with several possibilities: It could be that the HV lentiviral vector was not so mutagenic after all; alternatively the BAF3 cell line was not the most suitable target cell. This latter possibility was explored first.

A suitable target cell in mutagenesis experiments must allow selection for a phenotype, in this case IL-3 independence. It was thought a two to three day expansion of cells in the presence of IL-3 was sufficient to allow deregulation of cellular gene expression by proviral insertion. However, what seemed to happen was that upon IL-3 withdrawal, BAF3 cells died too rapidly before the selectable phenotype could take over in mutants. The role of the anti-apoptotic protein Bcl-2 expression in cancer cells is to prolong cell viability, allowing other mutations to accumulate. Bcl-2 over-expressing BCL15 cells remain viable for longer after IL-3 withdrawal and did allow IL-3 independent mutants to emerge, but again only after rescue with IL-3.

With the HV lentiviral vector, IL-3 independent mutants were obtained at a Q-PCR integrant frequency of 4.3×10^{-8} and cell frequency of 2.3×10^{-7} in experiments LV2 and LV3 combined. For mutants to grow out without rescue, the cell frequency needs to be at least 10^{-6} (See section 3.2.2.), meaning greater than 20 fold more integrants would have to be screened. This is not feasible, so in all other experiments a protocol that involved rescue with IL-3 was used.

4.3.3. Ghr gene activation in HV insertional mutants

The HV lentiviral vector is able to transform BCL15 cells to IL-3 independence. It seems to do so uniquely through activation of the Ghr gene. 15 out of 16 mutants obtained through standard selection express the Ghr transcript and independent HV vector integration sites in the Ghr locus were recovered in 10 mutants. As mentioned previously, the Ghr is a retroviral common insertion site, Evi124. The Ghr was a target of retroviral integration by endogenous Murine Leukaemia Virus (Akv) in 6 independent (B cell) leukaemias/lymphomas (Suzuki *et al.* 2002; Suzuki *et al.* 2006). In these two studies by Neal Copeland's group retroviral insertional mutagenesis was used with the objective of identifying new cancer associated genes. Ghr insertions in these tumours suggest that GH supports survival and/or proliferation of these cells in vivo. All insertions were clustered

within 5.5 kb up- or downstream of the L2 promoter, which drives transcription of the predominant Ghr transcript in mouse adult liver. Four out of 6 insertions were in reverse orientation as Ghr gene transcription, and would be expected to activate transcription from the L2 promoter by enhancer insertion.

All 10 identified insertions in my IL-3 independent mutants occurred in intron 1-2 in the same orientation as Ghr gene transcription. In contrast to wild type retrovirus insertions, none of the HV vector insertions were near the L2 promoter. All, except the insertion in mutant HV A2, were in the 15 kb region upstream of exon 2, near the L1 and L5 promoters. The most likely mechanism whereby HV vector insertion activated Ghr expression was considered to be enhancer activation by the HIV-1 or SFFV enhancers in the vector of one of the three endogenous promoters. However it was intriguing that all 10 identified HV vector insertions were in the same orientation as Ghr gene transcription. Generation of a read through transcript from the SFFV promoter was another possibility. This would necessitate inadequate polyadenylation at the 3'LTR.

4.3.4. HIV-1 promoter insertion results in the generation of a vector – Ghr fusion transcript

HIV-1 5' LTR activity in BAF3/BCL15 cells in the absence of the Tat transactivator protein was not anticipated. However, the HIV-1 major splice donor interacted with the splice acceptor of Ghr exon 2 to generate a vector-Ghr fusion transcript. Similar cases of promoter insertion were reported in wild type Mo-MLV induced rodent haematopoietic tumours. MLV insertions in the same orientation upstream of c-myb were found in mouse myeloid tumours (Shen-Ong *et al.* 1986). Transcripts were initiated from the 5'LTR and use a cryptic splice donor present in gag to splice to the splice acceptor of a downstream c-myb exon.

MLV promoter insertion activated expression of a cytokine receptor, namely the prolactin receptor (PRLR) in a rat T cell lymphoma (Barker *et al.* 1992). Proviral insertion occurred upstream of the Prlr gene. Transcripts were initiated from the proviral 5'LTR. The MLV splice donor upstream of the gag gene interacted with a cryptic splice acceptor in exon 1 of Prlr.

4.3.5. Why is the HIV-1 LTR active in BAF3/BCL15 cells?

In wild type HIV-1 infected cells basal LTR activity is required to re-activate viruses from latency. The transcription factors involved in basal transcription from the HIV-1 LTR include Sp1 (Jones *et al.* 1986) and NFκB (Nabel and Baltimore 1987). We hypothesise these transcription factors are active in BAF3 cells and initiate transcription from the HIV-1 LTR. Sp1 transcription factors are ubiquitously expressed, whereas NFκB is inducible upon antigen stimulation and cytokine signalling. The importance of these two transcription factors in HIV-1 LTR promoter activity in BAF3 cells was demonstrated when HIV-1 LTR luciferase reporter gene constructs were transfected into BAF3 cells expressing the human GM-CSF receptor (Watanabe *et al.* 2002). Maximal luciferase expression was seen in an HIV luciferase construct that contained the complete U3 region and the first 25bp of the R region. NFκB and Sp1 binding site deletion mutants displayed much reduced luciferase expression, highlighting the importance of these transcription factors in GM-CSF induced LTR activity.

The tissue specific transcription factor GATA-2 was also shown to be important in HIV-1 LTR promoter activity (Towatari *et al.* 1998). Luciferase expression in BAF3 cells from an HIV-1 LTR construct in which the GATA-2 binding sites were mutated was reduced by 60% compared to wild type. GATA-2 is phosphorylated by a MAPK-dependent pathway in BAF3 cells in response to IL-3 signalling (Towatari *et al.* 1995).

The role of the trans-activator protein Tat, encoded by wild type HIV-1, is in HIV-1 LTR transcriptional initiation and elongation. Tat is not expressed in BAF3 target cells and elongation of transcripts initiated from the HIV-1 5'LTR would expected to be very inefficient. A recent paper examined the role of NFκB activity in latent HIV-1 gene expression (Williams *et al.* 2007). NFκB is not only a strong inducer of HIV-1 transcriptional initiation; its sustained expression can induce transcriptional elongation. The RelA subunit of NFκB is thought to recruit P-TEFb, which phosphorylates serine 2 of the C-terminal domain of RNA polymerase II. In wild type HIV-1 infected cells these elongated transcripts code for small quantities of Tat, leading to efficient elongation of HIV-1 transcripts.

In the same way as GM-CSF, IL-3 signals through the common β chain and one would expect many of the same pathways to be activated by IL-3 and GM-CSF signalling in BAF3 cells. Following transduction, BAF3 cells are cultured in the presence of IL-3 for 2-3 days. IL-3 signalling at this time could activate HIV-1 LTR activity. IL-3 was shown to induce NFκB activity in BAF3 cells (Nakamura *et al.* 2002). Once the GHR is expressed on the cell surface of Ghr-insertional mutants, GH signalling itself induces NFκB activity (Jeay *et al.* 2000) and could in that way sustain its own expression. Future work will investigate the role of different transcription factors on HIV-1 LTR activity in BAF3 cells.

4.3.6. Why does HV vector transformation only work via activation of Ghr expression?

Assuming HIV-1 LTR activity is not just restricted to the Ghr locus, it should have the potential to activate other genes by promoter insertion. However, in this assay we only observed activation of the Ghr by this mechanism. The Ghr is unique in that it has a very large (125kb) first intron preceding the first coding exon. For this mechanism of promoter insertion to work, vector insertion needs to occur in the same orientation as gene transcription and there needs to be a splice acceptor upstream of the coding region that can interact with the splice donor of the vector. This implies that candidate target genes cannot have a coding first exon, unless there is a cryptic splice acceptor present upstream of the start codon in exon 1. This long list of prerequisites for this mechanism of gene activation to work, probably explains why we do not see activation of other candidate genes such as IL-3, Igf1, Pdgfr, Fgfr1 and Bcl-X, that could also potentially transform BAF3 cells to factor independence. IL-3 is a very small gene of just over 2 kb, whose first coding exon is exon 1. The Igf1 gene is much larger at 56kb, but again exon 1 is the first coding exon. Looking at the structure of genes encoding growth factor receptors whose ligands are present in serum, all transcripts contain at least one non-coding exon. The Pdgfr expresses 2 transcripts; translation starts at exon 2; the first introns are small at 457bp and 14.6 kb. The first coding exon of Fgfr1 is exon 2, but again the first intron is small at 13 kb. HIV-1 LTR promoter insertion might well activate expression of these

receptors, but their ligands might not be present at sufficient concentrations in serum or the resulting phenotype is not robust enough to survive the selection protocol. Retroviral vectors can activate Bcl-X_L expression (Thomas *et al.* 1998). The Bcl-X (Bcl2l1) gene is 64 kb in size. Four different transcripts are expressed from this gene. Exon 2 is the first coding exon in all of these transcripts. Intron 1-2 varies in size from 283 to 1500bp, so again is relatively small. BCL15 cells already over-express Bcl-2, so Bcl-X_L insertional activation would be unlike to affect the phenotype of such a cell. As discussed previously, Bcl-X_L over-expression alone in BAF3 cells does not result in IL-3 independent growth, but only prolongs viability. If we had many integrants per cell, we could select for Bcl-X insertions cooperating with insertions in a genes encoding mitogens. However, the MOI in these experiments was presumably too low for that. Finally, another reason why we might only pick up Ghr insertional mutants, is because the Ghr is a preferred integration site for HIV-1. Regional hotspots of HIV-1 integration have been reported (Schroder *et al.* 2002). However, in parental cells, the Ghr is not expressed and HIV-1 integrates preferentially but not uniquely in active genes (Schroder *et al.* 2002).

4.3.7. Why does the GH selection protocol work?

Ghr-insertional mutants are unable to expand when selected in serum alone without rescue with IL-3. However, in the final part of this chapter it was demonstrated that Ghr-insertional mutants could grow out when selected in serum supplemented with additional growth hormone.

It is difficult to estimate the bGH concentration in serum. Also, the issue of other factors present in serum that could synergise with bGH remains unresolved. It is thought very low levels of GH are sufficient to stimulate GH signalling. It could be that during the first 10 days after transduction, GHR-expression levels by Ghr-insertional mutants are still very low and this is why higher concentrations of bGH are needed during selection. This was demonstrated in BAF3 cells engineered to express the IL-2 receptor (Collins *et al.* 1994). The IL-2 concentration needed to support survival and proliferation of BAF3 clones expressing different levels of IL-2 receptor on the cell surface was compared. In clones that expressed low levels

of IL-2 receptor, high concentrations of IL-2 were needed, presumably to induce maximal receptor occupancy. As Ghr-insertional mutants become established, GHR expression levels might increase, allowing them to grow in serum without additional bGH.

Slightly more mutants (18 versus 13) were picked up after selecting the same transduced cell population in serum plus bGH, as opposed to standard selection. Interestingly, these were all independent clones. A total of 30 Ghr-insertional mutants were obtained after screening 1.9×10^8 integrants in experiment LV3. Overall, in mutagenesis experiments with the HV lentiviral vector in the BCL15 cell line 3.1×10^8 HV vector integrants were analysed. Assuming random integration, this means roughly one HV vector integrant per every 10bp of the genome. Intron 1-2 must have been hit 1.25×10^4 times and half of these integrants should be in the same orientation as Ghr gene transcription. This resulted in 33 Ghr insertional mutants, meaning approximately 1 in 200 integrants in the locus could result in a mutant.

Chapter 5

5. Insertional mutagenesis by a self-inactivating lentiviral vector

5.1. Introduction

In chapter 4, a frequency at which IL-3 independent mutants of the BCL15 cell line were obtained with the wild type HIV-1 LTR lentiviral vector HV, was established. The mechanism by which this vector transforms BCL15 cells was analysed. Insertional activation of the Ghr gene by the HV vector was found to have occurred in IL-3 independent mutants. A vector-Ghr fusion transcript originating from the HIV-1 5'LTR was found in these mutants. Activity of the HIV-1 LTR in the BCL15 cell line was unexpected.

Self-inactivating (SIN) lentiviral vectors should theoretically be less mutagenic in the BCL15 cell line, because they lack HIV-1 LTR promoter and enhancer activity.

In this chapter, the ability of the self-inactivating lentiviral vector CSGW (Demaison *et al.* 2002) to transform BCL15 cells to IL-3 independence was analysed. The SIN LTR in CSGW is derived from pHR SIN (Zufferey *et al.* 1998), described previously in section 1.3.4. of the introduction. It has a 400bp deletion of the U3 region and residual HIV-1 promoter activity by the 5'LTR should be negligible. The questions that were addressed were as follows:

- Is a SIN lentiviral vector, otherwise identical to the HV lentiviral vector used in chapter 4, still able to transform BCL15 cells to IL-3 independence and if yes, at what frequency?
- By what mechanism(s) does this SIN lentiviral vector transform BCL15 cells?

5.2. Results

5.2.1. The integrant frequency at which IL-3 independent mutants are obtained with a SIN lentivector is almost 10 fold lower than with a wild type LTR lentivector

Two experiments with the SIN lentiviral vector CSGW were performed. BCL15 cells were transduced with the CSGW lentiviral vector and selected according to the standard selection protocol, which involved rescue with IL-3 after an initial week of IL-3 starvation. In experiment LV4, 3 mutants were obtained in the vector transduced population versus one mock mutant (Table 5.1). This was at an integrant frequency of 9.8×10^{-9} by Q-PCR. The same transduced target cell population was also selected in serum plus bGH. In contrast to experiment LV3 with the HV lentiviral vector (Table 4.3.), this did not result in an increase in the number of vector transduced mutants that were obtained. More mutants were obtained from the mock than the vector transduced population (3 versus 2). In experiment LV5, 5.4×10^{8} CSGW integrants were screened; almost twice the number of integrants analysed in experiment LV4. No long term IL-3 independent mutants were obtained after either standard selection or selection in serum plus bGH. The difference between the observed cell frequencies in the mock and HV arms of the experiment in LV4 and LV5 combined was not significant in either the standard or serum plus bGH selections.

Table 5.1. Mutagenesis frequencies with the self-inactivating lentiviral vector CSGW

Experiment	Target Cell	Mock Mutants		CSGW Vector Transduced Mutants		
		Mutant Number	Cell Frequency [a]	Mutant Number	Cell Frequency [a]	Integrant Frequency [b]
Standard selection						
LV4	BCL15	1	2.0×10^{-8}	3	6.0×10^{-8}	3.3×10^{-8} 9.8×10^{-9}
LV5	BCL15	0	$< 2.0 \times 10^{-8}$	0	$< 2.0 \times 10^{-8}$	$< 1.0 \times 10^{-8}$ $< 1.9 \times 10^{-9}$
Selection in foetal calf serum plus 1 µg/ml bovine growth hormone						
LV4	BCL15	3	6.0×10^{-8}	2	4.0×10^{-8}	2.2×10^{-8} 6.5×10^{-9}
LV5	BCL15	0	$< 2.0 \times 10^{-8}$	0	$< 2.0 \times 10^{-8}$	$< 1.0 \times 10^{-8}$ $< 1.9 \times 10^{-9}$

Table 5.1. Mutagenesis frequencies with the self-inactivating lentiviral vector CSGW

In experiments LV4 and 5, 5 x 10^7 BCL15 target cells were transduced with the self-inactivating lentiviral vector CSGW. Cells were cultured in the presence of IL-3 for 48 hours following transduction. By day 3 of the experiment, target cells had expanded 8-10 fold in cell number. 10^8 cells were washed thrice in complete medium without IL-3.
5 x 10^7 cells each were selected in one of the ways, as described in Table 4.3.

[a] The cell frequency was calculated by dividing the number of mutants by the total number of target cells in the experiment.
The difference between the observed cell frequencies in the mock and HV arms of the experiment in LV4 and LV5 combined was significant in neither the standard nor FCS + bGH selections: standard selection $p = 0.0015$, $X^2 = 10$; FCS + bGH selection $p = 0.1572$; $X^2 = 2$.
[b] The integrant frequency was calculated by dividing the number of mutants by the total number of integrants in each experiment.
The number of "GFP" integrants was calculated by multiplying the number of live target cells by the MOI. The MOI was determined by flow cytometry for GFP expression at 48h to 72h post-transduction. The GFP integrant frequency is shown in green.
In experiments LV4 and LV5, the average number of vector copies per cell was also determined by Q-PCR. This number was multiplied by the number of target cells to obtain the number of "Q-PCR" integrants. The Q-PCR integrant frequency is shown in black.

5.2.2. The same two subclones seem to have been picked up following standard selection and selection in serum plus bGH

IL-3 independent mutants from experiment LV4 were analysed by Southern blot (Figure 5.1.) to determine clonality and vector copy number. Mutants carried 3 to 5 copies of the vector each. Analysis of the band pattern of standard selection mutants revealed that CSGW12 is the only unique mutant. The band patterns of CSGW91 and CSGW94 look similar; the top double band in CSGW91 is not very clear and there is a fifth band below the top double band. This would suggest CSGW91 and CSGW94 are subclones. Interestingly, the same two subclones seem to have been picked up by selection in serum plus bGH. The band patterns of bGH CSGW1-1 and bGH CSGW2-4 look similar to that of CSGW91 and CSGW94, respectively.

5.2.3. The SIN lentiviral vector mutants do not express the Ghr transcript, nor do they respond to bGH in culture

Before proceeding to analyse the integration sites in these mutants, I first assessed by RT PCR whether any of them were likely to be Ghr insertional mutants. None of the CSGW mutants produce the Ghr transcript (Figure 5.2.a). Because of the comparatively high number of mock mutants in experiment LV4 (4 versus 5 CSGW mutants), IL-3 expression levels were also analysed by Q-RT-PCR. All mock mutants and mutant CSGW12 express the IL-3 transcript (Figure 5.2.b). The two subclones CSGW91 (= GH CSGW 1-1) and CSGW94 (=GH CSGW2-4) express neither Ghr nor IL-3 transcript. This is a unique phenotype, I had not seen before. For this reason, I decided to focus my integration site analysis on these two mutants.

None of the CSGW mutants respond to bGH in culture (Figure 5.3.). Interestingly, no difference in the ability to grow in serum free medium can be seen between the mutants that express IL-3 (Mock60, CSGW12) and the others that do not.

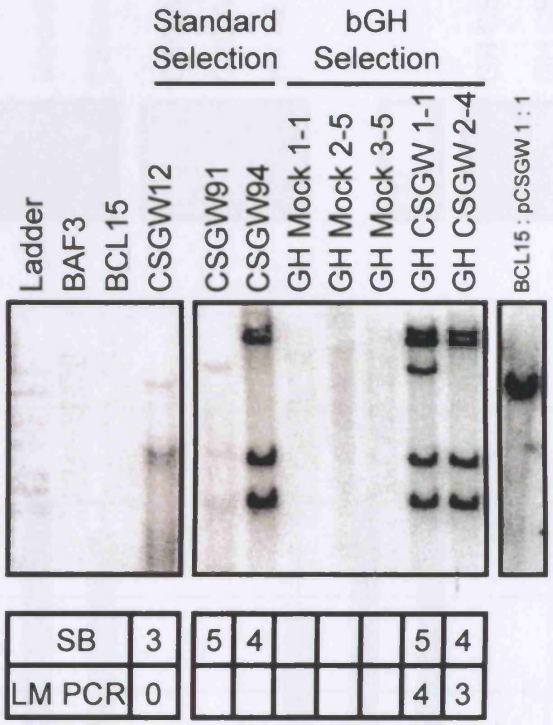


Figure 5.1. Southern blots of IL-3 independent mutants obtained in experiment LV4 following standard selection and selection in foetal calf serum supplemented with 1µg/ml bGH.

15µg genomic DNA of each mutant was digested with BamHI. This enzyme cuts only once inside the vector at the same position as in the HV vector (See Figure 4.2.). Blots were probed with the cDNA for eGFP. Blots from LV4 were exposed to a phosphorscreen. Irrelevant lanes on blots were removed.
Band patterns were examined to determine clonality of the mutants and CSGW vector copy number per genome. The number of bands seen on the Southern blots and the number of integrations sites cloned out by LM-PCR for each mutant are shown in the table below the blot.

The band pattern of CSGW12 is unique compared to the other two mutants picked up in the standard selection. There was less DNA in the CSGW91 lane. However, on closer inspection the band patterns of mutants CSGW91 and CSGW94 are similar, except that CSGW91 has a fifth band.
Furthermore, CSGW91 and GH CSGW1-1 show a similar band pattern on Southern blot, as do mutants CSGW94 and GH CSGW2-4. This suggests the same clones were picked up by standard selection and selection in FCS plus bGH.

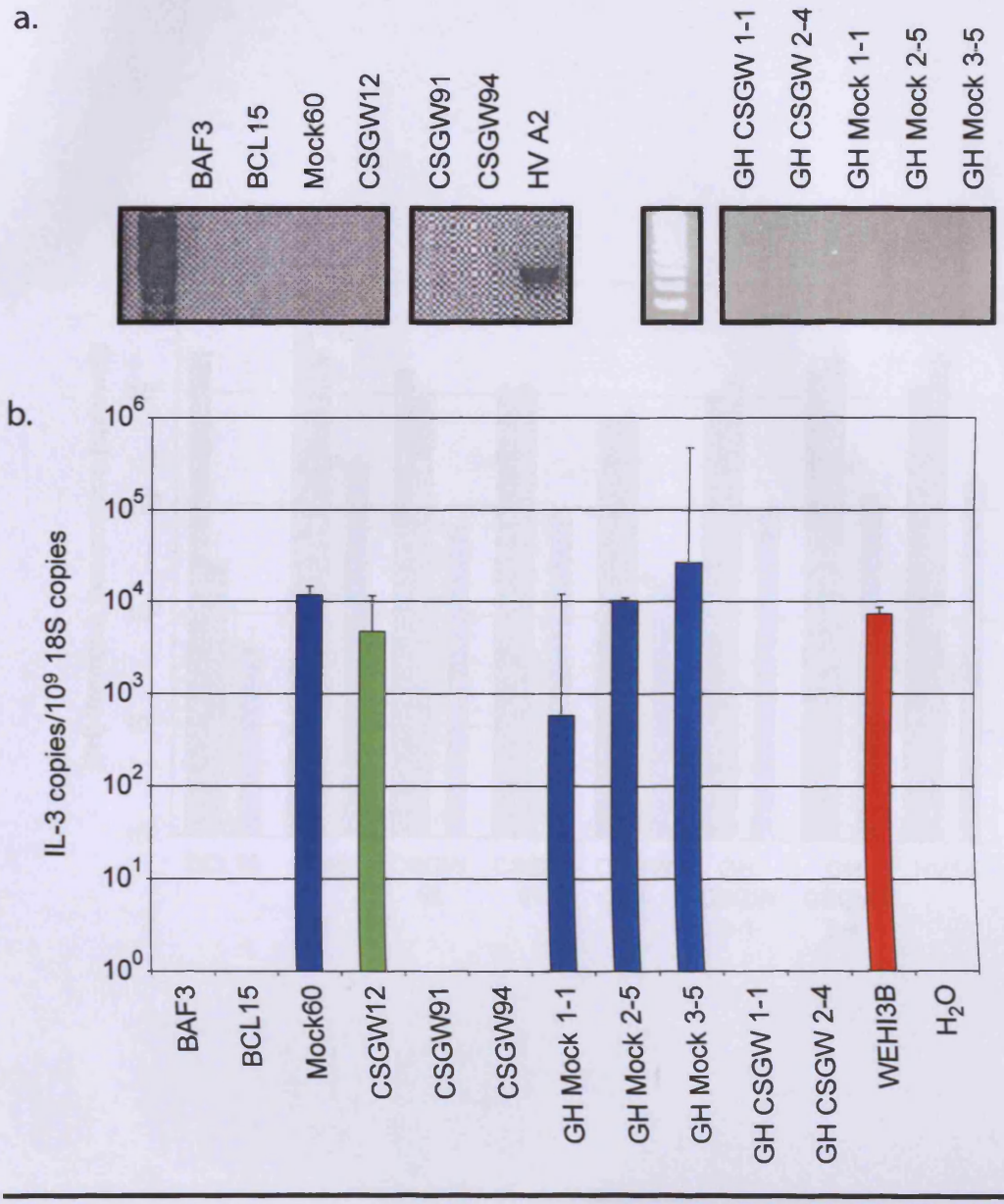


Figure 5.2. None of the CSGW mutants express the Ghr transcript.

a. A reverse-transcriptase (RT) PCR for expression of the Ghr transcript was performed on parental cells and IL-3 independent mutants from experiment LV4. A forward primer in exon 4 and a reverse primer in exon 8 were used to amplify a 552bp section of the GHR transcript. Mutant HV A2 RNA was used as a positive control. None of the CSGW mutants express the Ghr transcript.

b. Quantitative RT-PCR to measure IL-3 transcript levels was performed. Results are expressed as IL-3 copies per 10^9 18S copies. The error bars represent the standard error of this ratio. As positive control, RNA from WEHI3B cells was used. All four mock mutants as well as mutant CSGW12 express the IL-3 transcript.
Mutants CSGW91 and 94, and mutants GH CSGW1-1 and 2-4, express neither the Ghr nor IL-3 transcripts.

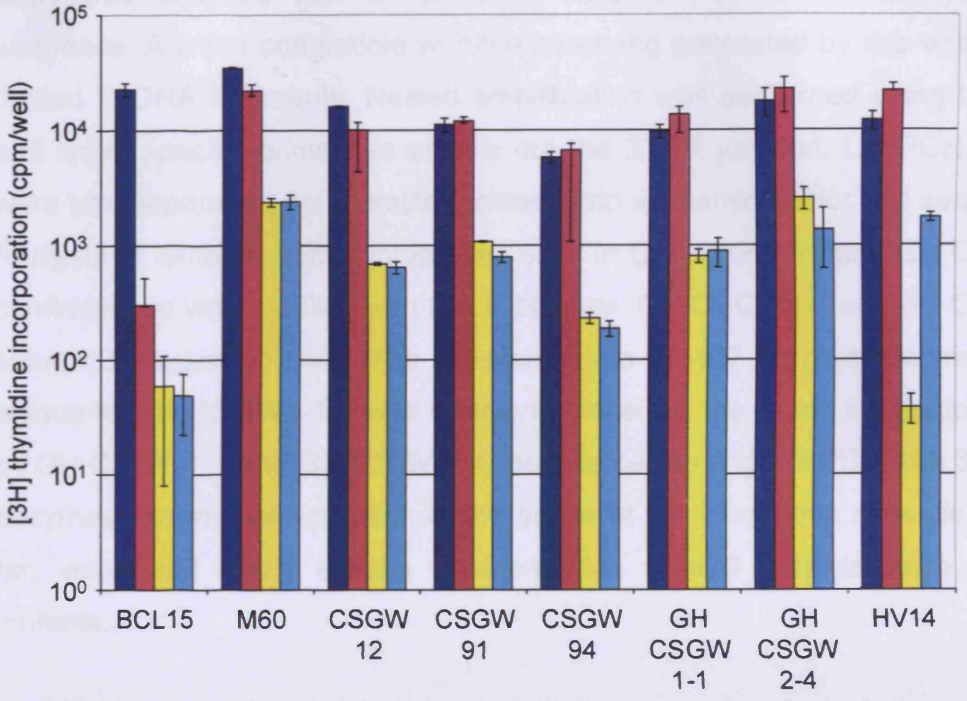


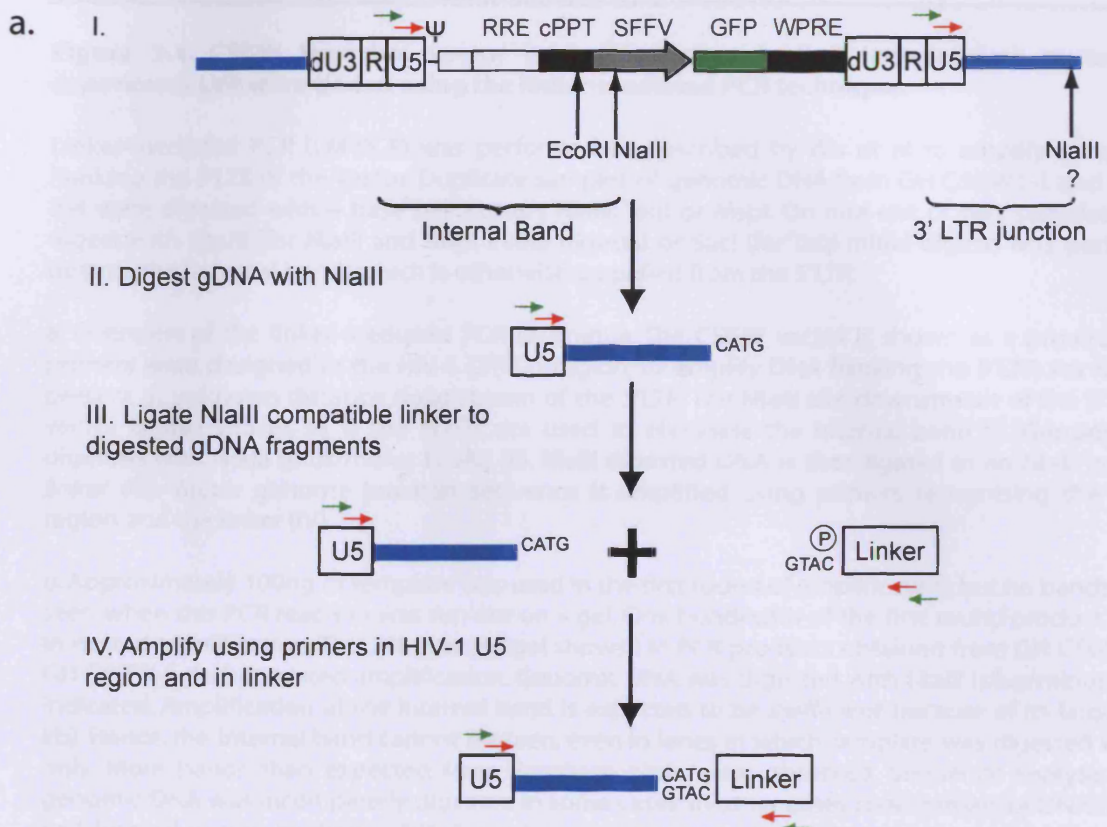
Figure 5.3. The mutants obtained with the CSGW lentiviral vector do not respond to bGH in culture.

Parental BCL15 cells and CSGW mutants were cultured in: medium containing 10% FCS and 10% WEHI (blue); medium containing 10% FCS only (pink); serum-free medium (yellow) or serum-free medium containing 1 µg/ml recombinant bGH (aqua blue).
Cells were cultured in these media for 24 hours before an overnight incubation in the presence of [3H] thymidine. [3H] thymidine incorporation was measured by liquid scintillation counting the following day.
Ghr insertional mutant HV14 was a positive control in this experiment. The three mock mutants obtained in the selection in serum plus bGH were not analysed. Error bars indicate the standard error of the mean calculated from 2 independent experiments.

5.2.4.Integration site analysis confirms two subclones were picked up, but does not fully reveal their genotype

SIN lentiviral vector integration sites were cloned out by linker-mediated PCR (LM PCR). The principle of this method is shown in Figure 5.4. In short, genomic DNA was digested with a restriction enzyme that has a 4bp recognition sequence. A linker compatible with the overhang generated by this enzyme was ligated to DNA fragments. Nested amplification was performed using U5 region and linker specific primers to amplify out the 3'LTR junction. LM PCR products were size separated, gel extracted, cloned into a plasmid vector and sequences.

Analysis of lentiviral vector integration sites in GH CSGW1-1 and GH CSGW2-4 confirmed we were dealing with two subclones. GH CSGW1-1 and GH CSGW2-4 shared 3 integration sites. The integration into Zfp407 on chromosome 18 was unique to GH CSGW1-1. I was unable to clone out the fourth integration shared by GH CSGW1-1 and GH CSGW2-4, despite performing LM PCR with 3 different enzymes. Insertional activation of the genes at the integration sites identified so far, would not easily explain transformation to IL-3 independence of these mutants.



C.

3'LTR Nested

GTCTGTTGTGTGACTCTGGTAAC →

GTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCACTTTACCACTC

ATTCAAGTATTACCTTTCACAACTCTAAACAAATGGATCTCACTCCCATTCTTTGACTACTGGTTTTCTGAATTCTGTTG

Linker Nested

← GTCCCTTAAGCGGAGCCCT

AGGGACATTTTGGCTGCTGTAGCAGATGTTTGAGACCCCATGTAGTCCCTTAAGCGGAGCCCT

Figure 5.4. a, b and c.

Figure 5.4. CSGW lentiviral vector integration sites in IL-3 independent mutants from experiment LV4 were cloned using the linker-mediated PCR technique.

Linker-mediated PCR (LM-PCR) was performed as described by Wu et al to amplify genomic DNA flanking the 3'LTR of the vector. Duplicate samples of genomic DNA from GH CSGW1-1 and GH CSGW 2-4 were digested with 4 base pair cutters NlaIII, TaqI or MspI. On one out of two samples a second digest with EcoRI (for NlaIII and MspI initial digests) or SacI (for TaqI initial digest) was performed, to destroy the "internal band", which is otherwise amplified from the 5'LTR.

a. Overview of the linker-mediated PCR technique. The CSGW vector is shown as a provirus. LM-PCR primers were designed in the HIV-1 LTR U5 region, to amplify DNA flanking the 3'LTR. An NlaIII site is present at unknown distance downstream of the 3'LTR. The NlaIII site downstream of the 5'LTR in the vector is also shown, as is the EcoRI site used to eliminate the internal band (I). Genomic DNA is digested with NlaIII (plus/minus EcoRI) (II). NlaIII digested DNA is then ligated to an NlaIII compatible linker (III). Vector genome junction sequence is amplified using primers recognising the HIV-1 U5 region and the linker (IV).

b. Approximately 100ng of template was used in the first round of amplification; but no bands could be seen when this PCR reaction was run out on a gel. One hundredth of the first round product was used in nested amplification. This 2% agarose gel shows LM PCR products obtained from GH CSGW1-1 and GH CSGW 2-4 after nested amplification. Genomic DNA was digested with NlaIII (plus/minus EcoRI) as indicated. Amplification of the internal band is expected to be inefficient because of its large size (1.8 kb). Hence, the internal band cannot be seen, even in lanes in which template was digested with NlaIII only. More bands than expected from Southern blots were observed. Sequence analysis revealed genomic DNA was incompletely digested in some cases; in other cases concatamers of DNA fragments had formed during ligation. LM-PCR products were gel extracted, cloned into a plasmid vector and sequenced.

c. The sequence of 223 bp LM PCR product obtained from GH CSGW1-1 and GH CSGW 2-4 is shown here. It contains 69 bp of HIV-1 U5 region (in green) and 21 bp of linker sequence (in black). The primers used in nested amplification are shown. The 133 bp of mouse genomic DNA sequence (in blue) mapped to the Cnot4 gene on chromosome 6. The NlaIII site which is reformed upon ligation of the linker is underlined.

Table 5.2. CSGW lentiviral vector integration sites in IL-3 independent mutants (LV4)

Clone	SB	LM-PCR	In Gene (Y/N)	Gene Name	Entrez Gene ID	Chr	Distance to TSS	Ori	Predicted Gene Function	RTCGD
LV4										
GH CSGW 1-1	5	4	Y	Zfp407	240476	18	23kb Intron 1-2	+		
			Y	Cnot4	53621	6	32kb Intron 1-2	-	CCR4 NOT Transcription Complex, subunit 4	
			Y	Gmeb2	229004	2	26kb Intron 4-5	-	Glucocorticoid modulatory element binding protein 2	
			Y	Wipi2	74781	5	12kb Intron 2-3	+	WD repeat domain, phosphoinositide interacting 2	
GH CSGW 2-4	4	3	Y	Cnot4	53621	6	32kb Intron 1-2	-	CCR4 NOT Transcription Complex, subunit 4	
			Y	Gmeb2	229004	2	26kb Intron 4-5	-	Glucocorticoid modulatory element binding protein 2	
			Y	Wipi2	74781	5	12kb Intron 2-3	+	WD repeat domain, phosphoinositide interacting 2	

Table 5.2. CSGW lentiviral vector integration sites in IL-3 independent mutants (LV4)

LM- PCR was performed on mutants GH CSGW1-1 and GH CSGW 2-4 using three different restriction enzymes, NlaIII, TaqI and MspI. PCR products were gel extracted, cloned into a plasmid vector and then sequenced. Vector-genome junction sequences were blasted against the mouse genome, released August 2007 (www.ensembl.org/Mus_musculus).

SB indicates vector copy number in the mutant determined by Southern blot. LM-PCR indicates the number of insertion site loci identified in the mutant by linker mediated PCR. All insertion sites shown in this table were confirmed by site-specific PCR. The distance to the transcription start site (TSS) of the nearest gene is shown. Ori indicates whether the vector insertion occurred in the same (+) or in the opposite (-) orientation as the gene's transcript. Predicted gene function is shown where known. None of the integration sites obtained from GH CSGW 1-1 and 2-4 were common insertion sites listed in the RTCGD.

5.3. Discussion

5.3.1. Summary of achievements in this chapter

In this chapter, it was firstly demonstrated that IL-3 independent mutants of the BCL15 cell line could still be obtained with a SIN lentiviral vector. The integrant frequency at which mutants were obtained was approximately 10 fold lower than with the HV vector. The difference in cell frequencies at which mutants were obtained from the mock and vector transduced populations was hardly significant. Secondly, it was shown that the SIN lentiviral vector transforms BCL15 cells by a different mechanism than the HV lentiviral vector. No Ghr insertional mutants were obtained with the SIN vector, which is consistent with the hypothesis that an active HIV-1 5'LTR promoter is required for this mechanism of transformation to work. Thirdly, the role of the vector in transformation of mutants to IL-3 independence could not convincingly be demonstrated.

5.3.2. Analysis of mutants obtained with a SIN lentiviral vector

Two IL-3 independent mutants that are subclones, were obtained with the CSGW lentiviral vector. These mutants express neither the Ghr nor the IL-3 transcripts, a phenotype that has not been seen previously. Presumably these two subclones expanded considerable during the 2 day bulk culture in the presence of IL-3, as they were both picked up following standard selection as well as selection in serum plus bGH. As these mutants do not express the GHR or respond to GH in culture, that means they were able to expand from amidst dead cells following unbiased selection, without the need for rescue with IL-3. This is unique as this type of mutant was never picked up before. Unfortunately, lentiviral integration site analysis of these mutants is incomplete, with one shared integration site still unidentified. The integration site might well hold the key to elucidating why these subclones are IL-3 independent. It is not obvious how activation of the genes found near the integration sites identified so far could result in IL-3 independence. There remains a possibility that these subclones are spontaneous mutants that happen to carry vector insertions, like mutant HV48 from experiment

LV3 in the last chapter. So far, we have only seen mock mutants that had spontaneously activated IL-3 expression. However, this need not be the only mechanism by which BCL15 cells can become spontaneously IL-3 independent. For example, point mutants in the IL-3 receptor, resulting in ligand independent signalling, could also do this.

5.3.3.Further analysis of which elements in the HV vector are required for Ghr insertional activation

The mutagenesis experiments described in this chapter show that SIN lentiviral vectors are unable to activate Ghr gene expression. This means HIV-1 promoter activity is effectively silenced in these vectors, and they should be safer to use in the clinical setting than wild type LTR lentiviral vectors.

We still do not know why the wild type HIV-1 LTR is active in BCL15 cells and able to drive transcription of elongated transcripts. For this reason, it would be interesting to perform a more detailed analysis of exactly which transcription factor binding sites in the HIV-1 U3 region are necessary for promoter activity. In chapter 4, it was speculated that NFκB, Sp1 and/or GATA-2 transcription factors might be important in HIV-1 LTR activity (Towatari *et al.* 1998; Watanabe *et al.* 2002) in BCL15 cells. HV vectors in which one of more binding sites for these transcription factors have been mutated could be tested for their ability to activate Ghr expression.

Transcription from the HIV-1 5'LTR in the Ghr insertional mutants is Tat-independent. It would be interesting to investigate whether the TAR loop is still required. It was recently reported that in a Tat-independent HIV-1 variant dependent on doxycycline for transcription, the TAR loop could be completely deleted without affecting transcription starting at the U3/R junction (Das *et al.* 2007).

Also, possible interactions between the internal elements within the HV vector cassette and the HIV-1 5'LTR would need to be investigated. An HV vector derivative lacking the Woodchuck post-transcriptional regulatory element (WPRE) is currently being tested, as is a construct in which the SFFV LTR promoter has been replaced by ubiquitously active cellular promoter.

The other obvious "mutagenic" component of the HV lentiviral vector is the HIV-1 major splice donor (MSD), because it allows interaction with the splice acceptor of Ghr exon 2. An obvious solution to further improve the safety profile of this vector would be to mutate the MSD. However, a functional MSD is required for inactivation of the poly(A) site in the 5'LTR (Ashe *et al.* 1995). The factor responsible for this was shown to be the U1 snRNP which interacts with the MSD, and in some way inhibits activity of poly(A) factors at the poly(A) site (Ashe *et al.* 1997). Mutating the MSD in lentiviral vectors would result in premature polyadenylation in the 5'LTR during vector production, resulting in low titre vector preparations.

Chapter 6

6. Retroviral vector mutagenesis in the BCL15 cell line

6.1. Introduction

In chapter 4, I demonstrated that a wild-type HIV-1 lentiviral vector could transform BCL15 cells to IL-3 independence via activation of Ghr gene expression. This was at an average integrant frequency of 4.3×10^{-8} by Q-PCR.

In this chapter, I assayed two retroviral vectors for insertional mutagenesis in the BCL15 cell line. The integrant frequencies at which mutants are obtained were compared to that of the HV lentiviral vector. I investigated whether the same host genes were targeted in mutants transformed by lentiviral versus retroviral vectors.

The two retroviral vectors tested were CNCG and MFG.S eGFP. The CNCG retroviral vector (Soneoka *et al.* 1995) was used in experiments RV1 and RV3. CNCG has two wild type Moloney MLV LTRs and contains an internal CMV promoter driving GFP expression. The MFG.S retroviral vector (Riviere *et al.* 1995) was tested in experiment RV2. This vector backbone was used in the X-SCID trials to deliver γc to target cells. The GFP reporter gene in this construct is expressed from the 5'LTR.

6.2. Results

6.2.1. Retroviral vector mutagenesis experiments in the BCL15 cell line

5 x 10^7 target cells each were transduced in RV1 and RV3. In the standard selection, 6 IL-3 independent mutants (C10, C40, C57, C96, C59 and C94) were obtained with the CNCG vector; versus zero in the mock transduced arm of the experiment (Table 6.1.). This was at a combined integrant frequency of 9.5 x 10^{-8} by Q-PCR and 9.0 x 10^{-8} by GFP. CNCG-transduced target cells in experiment RV3 were also selected in FCS plus 1µg/ml bovine GH. However, this method of selection did not result in any mutants. Ghr insertional activation is therefore not the main mechanism whereby CNCG transforms BCL15 cells.

One mutant, "G18", was obtained with the MFG.S vector in experiment RV2 at an integrant frequency of 1.8 x 10^{-7} by Q-PCR and 5.9 x 10^{-8} by GFP. The Q-PCR integrant frequency in RV2 is higher than the GFP integrant frequency. Surprisingly, the number of integrants determined by Q-PCR was 3 fold lower than that determined by GFP, whereas normally the reverse is seen. Single cell clones were cloned out of this MFG transduced population. Only 14% of the clones were GFP positive. This is consistent with the Q-PCR data which gave an average of 0.11 vector copies per cell in this experiment.

6.2.2. Characterisation of the retroviral vector mutants

The number of vector copies in these 7 retroviral vector mutants was determined by Southern blot (Figure 6.1.). On Southern blots of mutants C40, C96, C59 and G18 only a single band could be seen. Mutant C94 is most probably a spontaneous mutant: it does not show any bands on Southern blot. (It does not express GFP determined by flow cytometry; neither could integrated GFP copies be detected by Q-PCR (Data not shown)).

Mutants C10 and C57 seem to both contain several copies of the vector, but exactly how many could not be clearly made out. All mutants in RV1 and RV3 were independent clones.

Table 6.1. Retroviral vector mutagenesis frequencies after standard selection and selection in foetal calf serum plus 1 µg/ml bovine growth hormone

Exp	Vector	Target Cell	Mock Mutants		Retroviral Vector Transduced Mutants		
			Mutant Number	Cell Frequency [a]	Mutant Number	Cell Frequency [a]	Integrant Frequency [b]
Standard Selection							
RV1	CNCG	BCL15	0	$< 2.0 \times 10^{-8}$	4	8.0×10^{-8}	1.2×10^{-7} 7.0×10^{-8}
RV2	MFG.S	BCL15	0	$< 2.0 \times 10^{-8}$	1	2.0×10^{-8}	5.9×10^{-8} 1.8×10^{-7}
RV3	CNCG	BCL15	0	$< 2.0 \times 10^{-8}$	2	4.0×10^{-8}	5.9×10^{-8} 1.2×10^{-7}
Selection in foetal calf serum plus 1 µg/ml bovine growth hormone							
RV3	CNCG	BCL15	0	$< 2.0 \times 10^{-8}$	0	$< 2.0 \times 10^{-8}$	$< 2.9 \times 10^{-8}$ $< 5.9 \times 10^{-8}$

Table 6.1. Retroviral vector mutagenesis frequencies after standard selection and selection in foetal calf serum plus 1 µg/ml bovine growth hormone.

5 x 10^7 BCL15 target cells were transduced with the CNCG (RV1 and 3) or MFG.S eGFP (RV2) retroviral vectors. Cells were cultured in the presence of IL-3 for 48-72 hours following transduction. Cells were washed of IL-3 on day 3/4 of the experiment. In experiment RV1 and 2, 5 x 10^7 cells were selected only according to the standard selection protocol as described in Table 3.1. In experiment RV3, 5 x 10^7 cells each were subjected to either the standard selection protocol or selection in foetal calf serum plus 1µg/ml bGH, as described in Table 4.3.

[a] The cell frequency was calculated by dividing the number of mutants by the total number of target cells in the experiment. The average cell frequency at which mutants were obtained with CNCG was 6 x 10^{-8} (experiment RV1 and RV3). The difference between the observed cell frequencies in the mock and CNCG arms of the experiment was significant: RV1 and RV3 standard selection $p < 0.001$, $X^2 = 60$; RV2 standard selection $p < 0.001$; $X^2 = 20$.

[b] The integrant frequency was calculated by dividing the number of mutants by the total number of integrants in each experiment.
The number of "GFP" integrants was calculated by multiplying the number of live target cells by the MOI. The MOI was determined by flow cytometry for GFP expression at 48h to 72h post-transduction. The GFP integrant frequency is shown in green.
In experiments LV4 and LV5, the average number of vector copies per cell was also determined by Q-PCR. This number was multiplied by the number of target cells to obtain the number of "Q-PCR" integrants. The Q-PCR integrant frequency is shown in black.
The average integrant frequency at which mutants were obtained with CNCG was 9.0 x 10^{-8} (GFP) and 9.5 x 10^{-8} (Q-PCR).

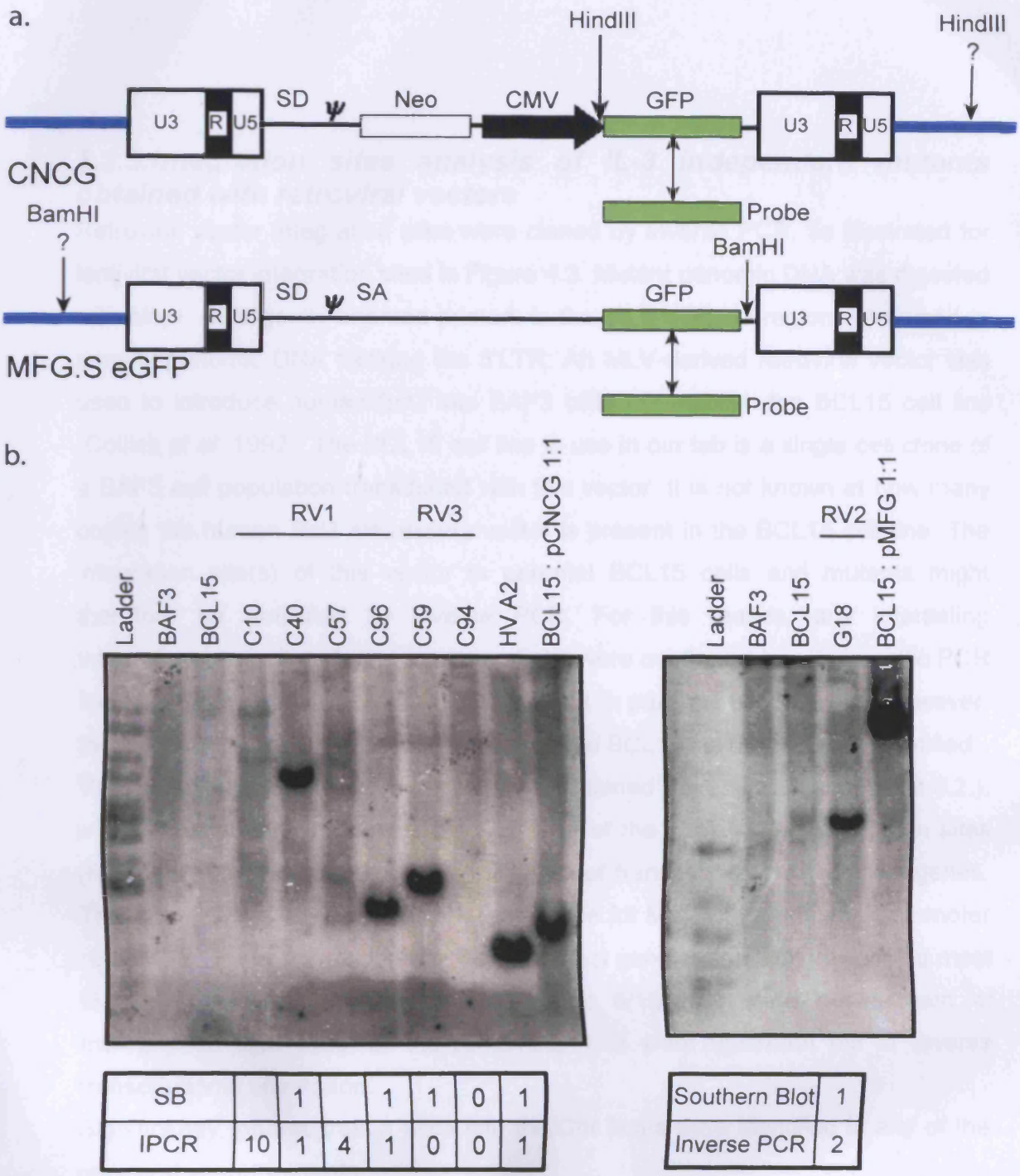


Figure 6.1. Southern blots of IL-3 independent mutants obtained in experiments RV1 – 3.

a. The retroviral vectors CNCG and MFG.S eGFP are shown as proviruses. The unique HindIII (CNCG) and BamHI (MFG.S eGFP) sites used in Southern blotting are indicated. Blots were probed with the cDNA for GFP.

b. 15µg genomic DNA of each mutant was digested with HindIII (RV1 and 3) or BamHI (RV2). Blots were exposed to autoradiography film. The tables below each blot show the estimated vector copy number and the number of integration sites in each mutant that were cloned out by inverse PCR.

6.2.3.Integration sites analysis of IL-3 independent mutants obtained with retroviral vectors

Retroviral vector integration sites were cloned by inverse PCR, as illustrated for lentiviral vector integration sites in Figure 4.3. Mutant genomic DNA was digested with MspI. Divergently oriented primers in the MLV LTR U3 region were used to amplify genomic DNA flanking the 5'LTR. An MLV-derived retroviral vector was used to introduce human Bcl2 into BAF3 cells to establish the BCL15 cell line (Collins *et al.* 1992). The BCL15 cell line in use in our lab is a single cell clone of a BAF3 cell population transduced with this vector. It is not known at how many copies the human Bcl2 expressing vector is present in the BCL15 cell line. The integration site(s) of this vector in parental BCL15 cells and mutants might therefore be amplified by inverse PCR. For this reason, any interesting integration sites cloned out by inverse PCR were confirmed by site-specific PCR for their presence only in the mutant and not in parental BCL15 cells. However, the integration site of the human Bcl2 vector in BCL15 cells was never identified.

Vector integration sites were successfully obtained from all mutants (Table 6.2.), with the exception of RV3 mutant C59. Out of the 18 retroviral integration sites (RIS) identified, 10 (56%) were within 5 kb of transcription start sites of genes. This is consistent with the reported preference for MLV integration near promoter regions. If these vector insertions were to affect gene expression, this would most likely be because of enhancer insertions: 6/10 RIS were downstream of transcription start sites, of the remaining 4, 3 were upstream but in reverse transcriptional orientation.

Significantly, no integration sites into the *Ghr* locus were identified in any of the retroviral vector mutants.

A RIS 2.6 kb downstream of the IL-3 gene transcription start site was identified in mutant G18. The GM-CSF (Csf2) gene transcription start site is located 17.4 kb downstream of the IL-3 gene; therefore, the MFG.S RIS was also 14.8 kb upstream of the GM-CSF gene in the same orientation as GM-CSF gene transcription.

In mutants C40 and C96, that each contained just a single copy of the vector, RIS near common insertion site (CIS) loci were identified. CIS are loci that were targets of wild type retrovirus integration in multiple independent mouse tumours and are likely to encode a cancer associated gene (Akagi *et al.* 2004).

In mutant C40, the CIS associated oxysterol binding protein like 3 (Osbpl3) gene was 46 kb downstream of the vector integration site. The RIS in mutant C96 occurred within the second intron of the Rps16 gene; the Evi24 CIS locus (Zfp36/ Plekhg2) is 20-30 kb downstream of this site. In mutant C57, four RIS were recovered. One of these was in semaphorin 4b (Sema4b), which is a CIS, Evi167. Furthermore, a RIS within a gene encoding a well known transcription factor GATA-1 was found. 10 RIS were recovered from mutant C10, but none of these were in CIS. The insertion site from mutant C59 was not recovered.

It is not immediately evident by which mechanisms, insertional dysregulation of the genes near the RIS in the CNCG mutants results in IL-3 independence of these mutants. The possible significance of insertions near CIS loci will be discussed later.

Table 6.2. Retroviral vector integration sites in IL-3 independent mutants (RV1-2)

Clone	SB	IPCR	In Gene (Y/N)	Gene Name	Entrez Gene ID	Chr	Distance to TSS	Ori	Predicted Gene Function	RTCGD
RV1										
C10		10	Y	Slc36a3	215332	11	3539bp intron 3-4	+	Proton amino acid transporter	
			Y	Mamdc2	71738	19	96kb intron 8-9	-		
			N			16				
			N			4				
			Y	Psma6	26443	12	685bp intron 1-2	+	Endopeptidase/hydrolase activity	
			N			13				
			N	Zdhhc4	72881	5	165bp upstream	-	Zinc ion binding; acetyltransferase activity	
			N			4				
			N	Mvd	192156	8	1030bp upstream	-	Diphosphomevalo-nate decarboxylase activity	
			N	Dusp6	67603	10	4740bp downstream		Dual specificity phosphatase 6	
C40	1	1	N	Osbpl3	71720	6*	46kb upstream	+	Lipid transport; steroid metabolism	2
C57		4	Y	Myl4	17896	11*	802bp intron 1-2	+	Motor activity	
			N	Cdc27	217232		1067bp upstream	-	Cell division	
			Y	Sema4b	20352	7*	21kb Intron 2-3	+	Cell differentiation; nervous system development	*Evi167* 5
			N			9*				
			Y	Gata1	14460	X*	4241bp Intron 1-2	+	Transcription factor activity	
C96	1	1	Y	Rps16	20055	7	779bp Intron 2-3	+	Structural constituent of ribosome	
			N	Zfp36/ Plekhg2	101497		21kb downstream	-	Guanyl nucleotide exchange factor activity	*Evi24* 9

Clone	SB	IPCR	In Gene (Y/N)	Gene Name	Entrez Gene ID	Chr	Distance to TSS	Ori	Predicted Gene Function	RTCGD
RV2										
G18	1	2	N	Il3	16187	11*	2607bp downstream	+	Cytokine/growth factor activity	
			N	Fbxo9	71538	9	3183bp upstream	+	Ubiquitin-protein ligase activity	
			N	Ick	56542		5744bp upstream	-	Protein kinase activity; signal transduction	

Table 6.2. Retroviral vector integration sites in IL-3 independent mutants (RV1-3)

Inverse PCR was performed on IL-3 independent mutants from RV1-2 using the restriction enzyme MspI. PCR products were gel extacted, cloned into a plasmid vector and then sequenced. Vector-genome junction sequences were blasted against the mouse genome, released August 2007 (www.ensembl.org/Mus_musculus). No RIS were obtained from RV3 mutants C59 and C94.

SB indicates vector copy number in the mutant estimated by Southern blot. IPCR indicates the number of insertion site loci identified in the mutant by inverse PCR. The asterisks in the chromosome column indicate that the integration site found in this locus was verified by site-specific PCR. Site-specific PCR was not performed on any of the integration sites found in mutant C10. The integration site in mutant C96 and the integration site in chromosome 9 in mutant G18 could not be confirmed by site-specific PCR.

The distance to the transcription start site (TSS) of the nearest gene is shown. Ori indicates whether the vector insertion occurred in the same (+) or in the opposite (-) orientation as the gene's transcript. Predicted gene function is shown where known. The number of appearances in the retroviral tagged cancer gene database (RTCGD) of each insertion site locus is shown where applicable.

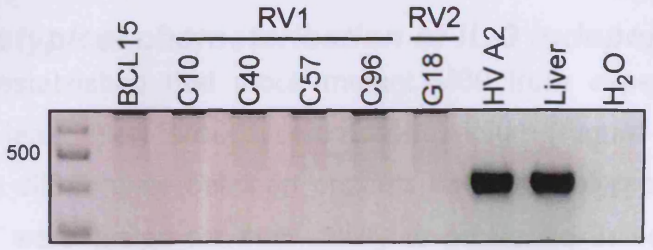
6.2.4.None of the IL-3 independent retroviral vector mutants express the Ghr transcript; all except one express the IL-3 transcript

To confirm that retroviral insertional mutagenesis did not occur through activation of the Ghr gene, RT-PCR for the Ghr transcript was performed. None of the 7 mutants obtained with CNCG or MFG.S expressed the Ghr transcript (Figure 6.2.a). Mutant G18 would be expected to express the IL-3 transcript, as it has a RIS downstream of the IL-3 gene. A normal size transcript encompassing the complete coding region could be detected by RT-PCR (Data not shown).

Earlier experiments suggested that activation of IL-3 gene expression was a mechanism whereby BCL15 cells could spontaneously become IL-3 independent. IL-3-Q-RT-PCR was therefore performed on all mutants by Mustafa Ceylan, an MSc student in the lab (Figure 6.2.b). Only mutant C59 did not express the IL-3 transcript. Mutant C94, which does not contain any copies of the vector, is a mock mutant that has spontaneously activated IL-3 expression. Mutant G18, which has a RIS downstream of IL-3, expresses the IL-3 transcript at very similar levels to C94. All mutants from experiment RV1 also express the IL-3 transcript, but at somewhat lower levels than G18 and C94. These mutants have either spontaneously activated IL-3 gene expression or insertional activation of genes near the retroviral integration sites resulted in IL-3 gene expression.

Mustafa Ceylan also measured expression levels of genes near RIS in mutants C40 (Osbpl3), C57 (Sema4b and Gata1) and C96 (Plekhg2 and Zfp36). He found no difference in expression levels of these genes between parental BCL15 cells and mutants (Data not shown).

a.



b.

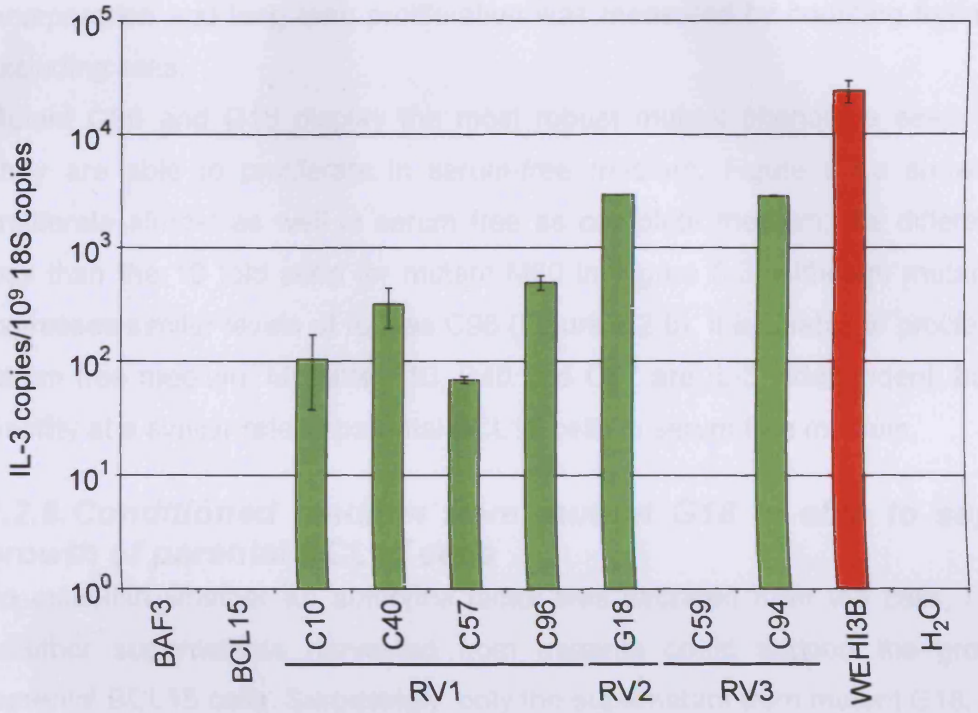


Figure 6.2. The IL-3 independent mutants obtained with retroviral vectors in experiments RV1-3 do not express the Ghr transcript; all but one of them express the IL-3 transcript.

a. A reverse-transcriptase (RT) PCR for expression of the Ghr transcript was performed on parental cells and IL-3 independent mutants from experiments RV1-3. A forward primer in exon 4 and a reverse primer in exon 8 were used to amplify a 552bp section of the GHR transcript. Mouse liver and mutant HV A2 RNA was used as a positive control. None of the mutants obtained with retroviral vectors express the Ghr transcript. Data for the two RV3 mutants is not shown.

b. Quantitative RT-PCR to measure IL-3 transcript levels was performed. Results are expressed as IL-3 copies per 10^9 18S copies. The error bars represent the standard error of this ratio. As positive control, RNA from WEHI3B cells was used. Only mutant C59 from experiment RV3 does not produce the IL-3 transcript. Mutant C94 from experiment RV3 appears to be a spontaneous mutant, as it was established by Southern blot and Q-PCR that it does not contain any copies of the vector. C94 expresses the IL-3 transcript at similar levels to mutant G18, which has an MFG vector integration site just downstream of the IL-3 gene. All mutants obtained in experiment RV1 also express the IL-3 transcript.

6.2.5. Phenotypical characterisation of IL-3 independent mutants

I previously established that mock mutant M60 from experiment LV4 could proliferate, at least short term, in serum free medium (Figure 5.3. in chapter 5). To be able to differentiate between mutants on a phenotypical basis, RV1 and RV2 mutants were tested for their ability to proliferate in serum-free medium (Figure 6.3). Short term proliferation was measured by [3H] thymidine incorporation and long term proliferation was measured by counting trypan blue excluding cells.

Mutant C96 and G18 display the most robust mutant phenotype seen so far. They are able to proliferate in serum-free medium. Figure 6.3.a shows they proliferate almost as well in serum free as complete medium; the difference is less than the 10 fold seen for mutant M60 in Figure 5.3. Although mutant C40 expresses similar levels of IL-3 as C96 (Figure 6.2.b), it is unable to proliferate in serum free medium. Mutants C10, C40 and C57 are IL-3 independent, but lose viability at a similar rate to parental BCL15 cells in serum free medium.

6.2.6. Conditioned medium from mutant G18 is able to support growth of parental BCL15 cells

To establish whether an autocrine factor was secreted from the cells, I tested whether supernatants harvested from mutants could support the growth of parental BCL15 cells. Surprisingly, only the supernatant from mutant G18, but not C96, was able to support growth of parental BCL15 cells (Figure 6.4.). Supernatant from mutant C96 was as ineffective as supernatant from mutant HV A2, which acted as a negative control. Supernatant from mutants C10, C40 and C57 supported growth of parental BCL15 somewhat at a concentration of 50%. The readings are lower at 100% for all supernatants; this is because at this data point cells were re-suspended in neat supernatant. As the cells were quite dense when the supernatant was harvested, it presumably contained some inhibitory factors that were diluted out at other data points.

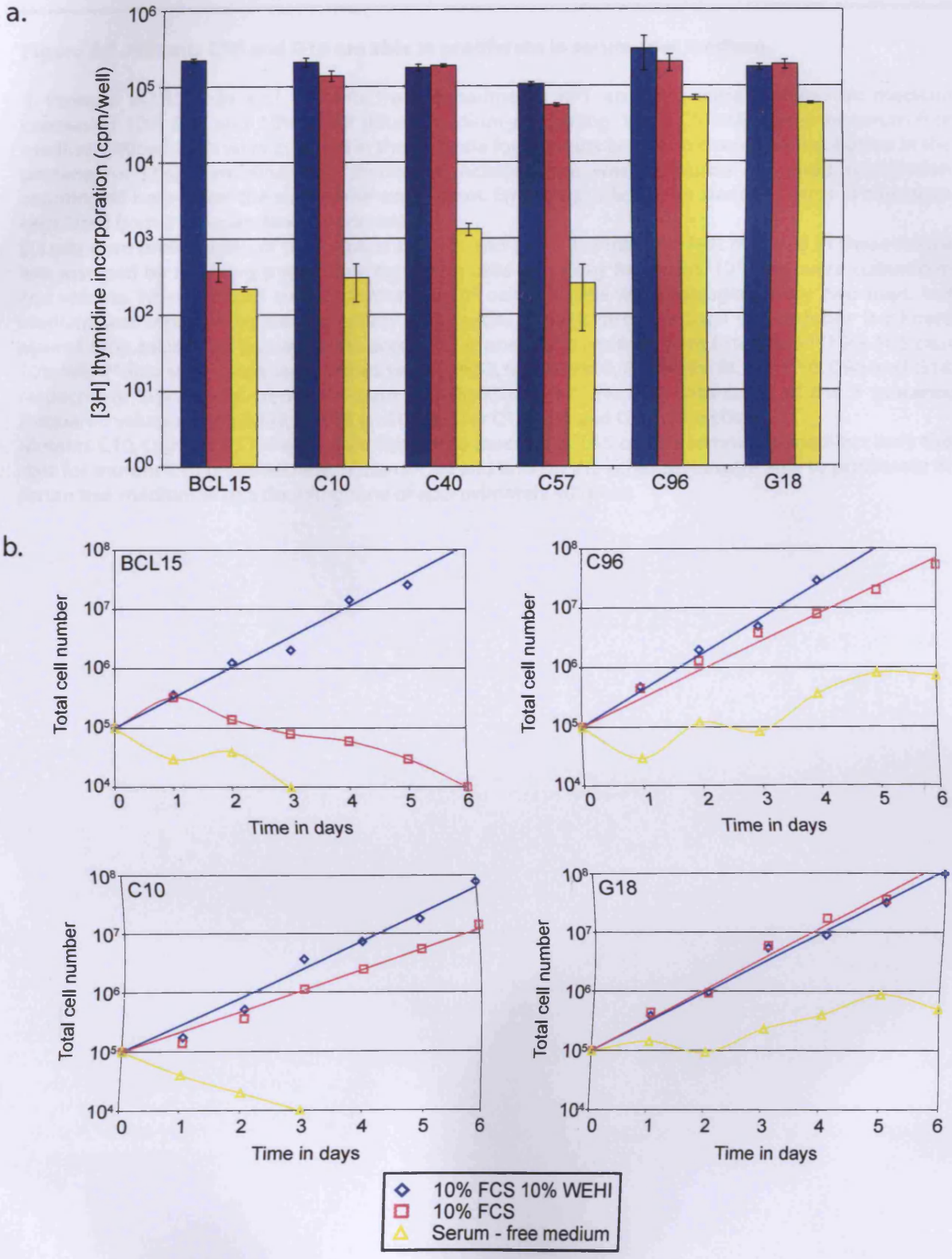

a.
[3H] thymidine incorporation (cpm/well)
BCL15
C10
C40
C57
C96
G18
b.
BCL15
C96
C10
G18
Total cell number
Time in days
10% FCS 10% WEHI
10% FCS
Serum - free medium

Figure 6.3.

Figure 6.3. Mutants C96 and G18 are able to proliferate in serum-free medium.

a. Parental BCL15 cells and mutants from experiments RV1 and RV2 were cultured in: medium containing 10% FCS and 10% WEHI (blue); medium containing 10% FCS only (pink) or serum-free medium (yellow). Cells were cultured in these media for 24 hours before an overnight incubation in the presence of [3H] thymidine. [3H] thymidine incorporation was measured by liquid scintillation counting 48 hours after the start of the experiment. Error bars indicate the standard error of the mean calculated from 2 independent experiments.
b. Long term proliferation of BCL15 parental cells and the 5 IL-3 independent mutants in these media was assessed by counting trypan blue excluding cells every day for 6 days. 10^5 cells were cultured in 2ml volume. When the cell count approached 10^6 cells/ml, cells were passaged. Every two days, 1ml medium was replaced by fresh medium in all wells. In the charts, the total cell number is plotted against time, taking cell passages into account. Exponential trendlines were fitted to all "10% FCS plus 10% WEHI" data series. R-squared values were 0.9852, 0.98, 0.9809, 0.9896 for BCL15, C10, C96 and G18, respectively. Exponential trendlines were also fitted to the "10% FCS" data series of the 3 mutants; R-squared values were 0.9839, 0.9933 and 0.9906 for C10, C96 and G18, respectively.
Mutants C10, C40 and C57 died at a similar rate to parental BCL15 cells in serum free medium; only the data for mutant C10 is shown here. Mutants C96 and G18 on the other hand were able to proliferate in serum free medium with a doubling time of approximately 40 hours.

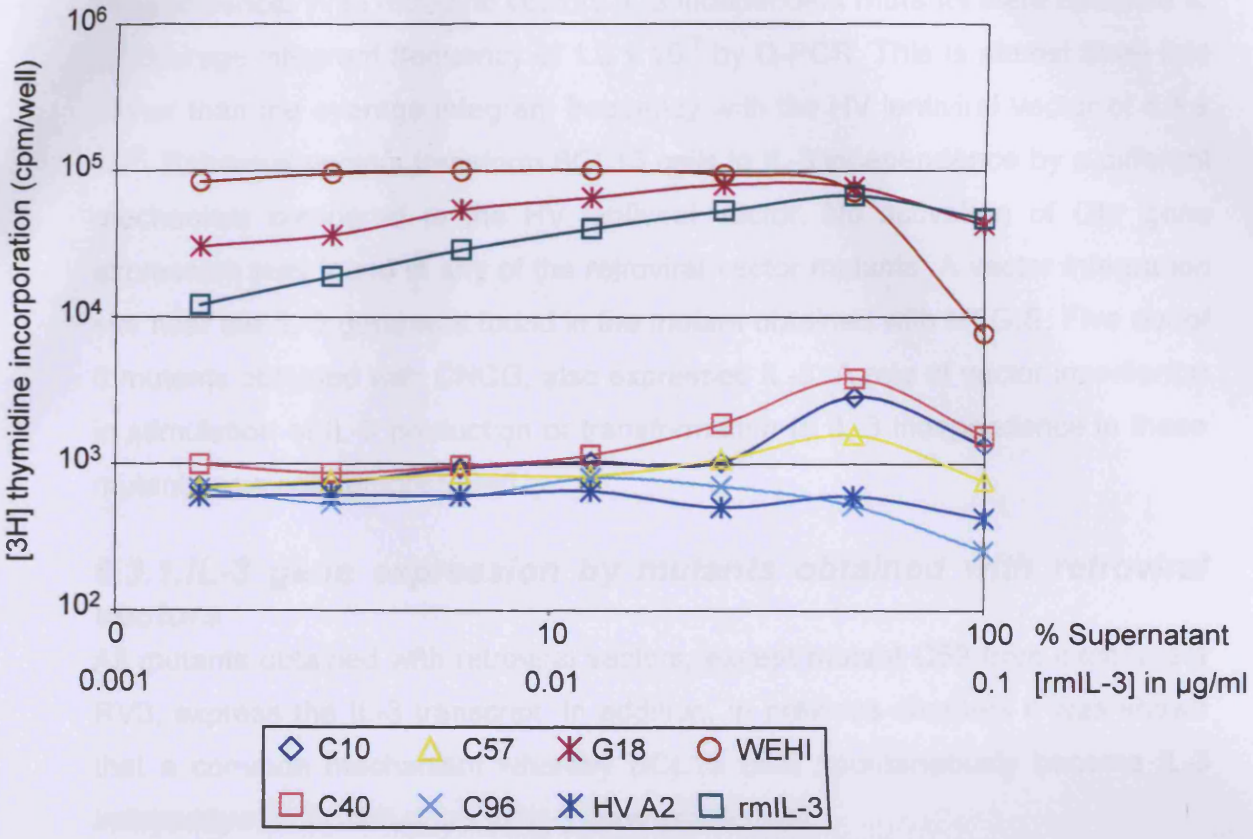


Figure 6.4. Supernatant from mutant G18 only is able to support growth of parental BCL15 cells.

Supernatants were collected from mutants when they were at a cell density of 10^6 cells per ml. Two fold serial dilutions of supernatant, starting at 100%, were plated out on 2 x 10^4 cells in a total volume of 200µl. Cells were cultured in these conditioned media for 24 hours before an overnight incubation in the presence of [3H] thymidine. [3H] thymidine incorporation was measured by liquid scintillation counting 48 hours after the start of the experiment.
Supernatant from WEHI3B cells acted as a positive control. Two fold serial dilutions of recombinant mouse IL-3, starting at a concentration of 0.1 µg/ml, served as a second positive control. Supernatant from mutant HV A2 served as a negative control. The mean value of triplicate wells is shown.

6.3. Discussion

The two retroviral vectors tested can transform BCL15 cells to IL-3 independence. With retroviral vectors IL-3 independent mutants were obtained at an average integrant frequency of 1.2 x 10^{-7} by Q-PCR. This is almost three fold higher than the average integrant frequency with the HV lentiviral vector of 4.3 x 10^{-8}. Retroviral vectors transform BCL15 cells to IL-3 independence by a different mechanism compared to the HV lentiviral vector. No activation of Ghr gene expression was found in any of the retroviral vector mutants. A vector integration site near the IL-3 gene was found in the mutant obtained with MFG.S. Five out of 6 mutants obtained with CNCG, also expressed IL-3. A role of vector insertion(s) in stimulation of IL-3 production or transformation to IL-3 independence in these mutants was not demonstrated.

6.3.1.IL-3 gene expression by mutants obtained with retroviral vectors

All mutants obtained with retroviral vectors, except mutant C59 from experiment RV3, express the IL-3 transcript. In addition, in previous chapters it was shown that a common mechanism whereby BCL15 cells spontaneously become IL-3 independent is by activation of IL-3 gene expression.

At this point, it would be useful to briefly discuss regulation of murine IL-3 gene expression, on a transcriptional and post-transcriptional level. I shall then proceed to consider the implications of this on IL-3 gene expression by the mutants.

6.3.1.1. Regulation of murine IL-3 gene expression

As mentioned in the introduction to chapter 3, IL-3 is normally only expressed by activated T cells, NK cells, mast cells and some megakaryocytic cells (Blalock *et al.* 1999). The murine IL-3 promoter is considerably less well studied than the human IL-3 promoter. The two promoters share 59% sequence homology (Osborne *et al.* 1995). Cytokine (CK) 1 and 2 elements, GATA, AP-1 and core binding factor (CBF) binding sites are found in both promoters. The TATA box of the mouse IL-3 gene is located 30bp upstream of the transcription start site. A

critical transcription factor involved in the regulation of IL-3 gene expression is Nuclear Factor of Activated T cells (NFAT) (Osborne *et al.* 1995; Ko *et al.* 2007). NFAT is activated following TCR signalling in T cells and Fcε crosslinking on mast cells. Phosphorylated NFAT is localised in the cytoplasm; upon calcium signalling it is dephosphorylated by calcineurin and subsequently translocates to the nucleus. Cyclosporin inhibits calcineurin and NFAT mediated IL-3 gene expression was found to be cyclosporin sensitive. In the mouse IL-3 gene locus three NFAT binding regions were identified. The first one is in the IL-3 promoter region (between -293 and -150bp); the second one is in intron 1-2 of the IL-3 gene (Ko *et al.* 2007). NFAT binding sites were also identified in an enhancer located 11kb downstream of the IL-3 transcription start site (Osborne *et al.* 1995). This enhancer activates transcription of both the IL-3 and GM-CSF genes. The latter is located 14kb downstream of the IL-3 gene on chromosome 11.

Post-transcriptional regulation of IL-3 gene expression is regulated by an AU-rich element (ARE) in the 3'UTR of the IL-3 transcript (Gillis and Malter 1991). This element binds hnRNP-D, which destabilises IL-3 mRNA, resulting in rapid turnover of IL-3 transcripts.

In the IL-3 expressing mutants C10, C40, C57 and C96 that do not contain vector integration site near the IL-3 gene, we could measure the levels of transcription factors regulating IL-3 gene expression.

6.3.1.2. Activation of IL-3 gene expression by transposable elements and retroviral vectors

In the WEHI3B mouse myelomonocytic cell line, an IAP insertion upstream of the IL-3 gene resulted in constitutive IL-3 expression by the cells. IAP insertion occurred in reverse orientation with respect to IL-3 transcription, so IL-3 activation is a result of enhancer insertion (Ymer *et al.* 1985; Ymer *et al.* 1986). Disruption of the 3'UTR ARE by IAP insertion results in increased IL-3 mRNA stability as well as IAP enhancer mediated activation of gene expression (Mayo *et al.* 1995).

6.3.1.3. Mechanism of IL-3 gene activation in mutant G18

In the G18 mutant from experiment RV2, the MFG.S retroviral vector insertion site (RIS) occurred downstream of the IL-3 gene. The insertion occurred 2.6kb downstream of the IL-3 transcription start site and did not disrupt the 3'UTR ARE. Since an endogenous enhancer is located 11kb downstream of the transcription start site, it is not surprising that a retroviral vector 2.6 kb downstream is able to activate transcription. This vector insertion also has the potential to activate GM-CSF gene expression by an enhancer mechanism. However, I was unable to demonstrate this by RT-PCR.

6.3.1.4. Possible mechanisms of IL-3 gene activation in the other four retroviral mutants

In four other mutants obtained with the CNCG retroviral vector, IL-3 gene expression was also activated, yet no RIS in or near the IL-3 gene were identified in the mutants. Several possibilities exist by which these mutants activated IL-3 gene expression to become IL-3 independent.

Firstly, I explored whether activation of CIS genes or other genes near RIS could result in activation of IL-3 gene expression. The following CIS insertions were found in mutants C40, C57 and C96:

Mutant C40

The integration site in mutant C40 is intergenic. The nearest transcription start site is 46 kb downstream, that of the Osbpl3 (oxysterol binding protein like 3) gene. The Osbpl3 locus is listed as a CIS in the RTCGD database. In a mouse model of pre-B cell leukaemia, the rate of which could be accelerated by Mo-MLV induced insertional mutagenesis, two RIS sites in this locus were identified (Bijl *et al.* 2005). Genes that collaborated with the E2a-PBX1 fusion protein were identified using this approach. 18% of integrations in these pre B cell leukaemias were found in the Hoxa locus on chromosome 6. The Hoxa locus, itself a CIS (Evi7), encodes homeobox transcription factors expressed in haematopoietic cells. The Hoxa locus is 2 Mb upstream of Osbpl3. Instead of representing a new proto-oncogene, RIS in the Osbpl3 locus are thought to also act via Hoxa gene

up-regulation in pre-B cell leukaemias. We have not looked at expression levels of Hoxa genes in mutant C40.

Mutant C96

The integration site in mutant C96 is in a gene dense region. It is located in intron 2-3 of Rsp16s (ribosomal protein 16S). However, this RIS is also 21/ 28 kb downstream of the Plekhg2/ Zfp36 transcription start sites. Plekhg2, also known as Clg, encodes a guanine nucleotide exchange factor and is a CIS, Evi24 (Himmel *et al.* 2002). Insertions upstream of this gene activate Plekhg2 expression in B cell and myeloid leukaemias.

Mutant C57

In mutant C57, four RIS were found. One of these was in semaphorin 4b (Sema4b), a CIS (Evi167). The insertion was in the same transcriptional orientation in the second intron of Sema4b. Out of the 5 RIS found in mouse tumours in this locus, 3 occurred upstream of the gene; the other ones were in intron 1 and exon 2, which is the first coding exon of this gene.

The RIS in Gata1, though not a CIS, is also potentially interesting as it occurred upstream of the first coding exon 2 in the same transcriptional orientation. Upon IL-3 signalling in BAF3 cells, the transcription factor GATA-1 is phosphorylated on serine 26 by a MAPK-dependent pathway. It is then mediates transcription of the survival gene E4bp4, which is itself a transcription factor, and Bcl-X_L (Yu *et al.* 2002; Yu *et al.* 2005). The same group showed that BAF3 over-expressing GATA-1 maintained cell viability for longer upon IL-3 withdrawal (Yu *et al.* 2002).

It was hypothesised that CIS gene activation could somehow up-regulate activity of transcription factors controlling IL-3 gene expression. However, we were unable to demonstrate CIS gene up-regulation in these mutants when cultured in the absence of IL-3; neither could up-regulation of GATA-1 transcription factor expression be demonstrated in mutant C57. It remains possible that retroviral vector insertion alters CIS gene expression for example on initial IL-3 removal.

It would be interesting to analyse RIS in a few random, IL-3 dependent clones, that were transduced with retroviral vectors, but were not selected for IL-3 independence. If RIS in CIS loci were found in these random clones, that would suggest they were probably not significant in transformation to IL-3 independence of our mutants and the recovery of RIS in CIS in them was just an incidental finding. It is also possible that retroviral vector insertion could have affected expression of genes much further a field than we investigated. These genes could be responsible for the activation of IL-3 expression. Large scale analysis of gene expression in these mutants by microarray would have to be performed to identify such genes.

There is no data to suggest that either Hoxa, Plekhg2 or GATA-1 gene up-regulation will activate IL-3 gene expression. However, up-regulation of expression of these genes in haematopoietic cells in combination with IL-3 is possible. In this model, insertional activation of genes near CIS is not responsible for activation of IL-3 gene expression per se, but cooperates with IL-3 in cells that have switched on low level expression of this cytokine, to produce an IL-3 independent phenotype.

Lastly, retroviral vector insertion might not have been involved in transformation to IL-3 independence of these mutants at all. In previous experiments, it was observed that activation of IL-3 gene expression was a common mechanism by which BCL15 cells spontaneously became IL-3 independent. The mechanisms by which this can occur will be discussed below.

6.3.1.5. Mechanisms of spontaneous activation of IL-3 gene expression

In the introduction to chapter 3, a study was cited in which mutants of the IL-3 dependent cell line FDC-P1(M) were selected for activation of GM-CSF alpha chain expression following retroviral vector transduction (Laker *et al.* 2000) (See Table 3.1). In 40% of the GM-CSF alpha chain expressing mutants analysed, gross chromosomal rearrangements of the GM-CSF alpha chain locus were found to have occurred. It is conceivable that this mechanism could also be responsible for spontaneous transformation to IL-3 independence of BCL15 cells. We have not looked for deletions, insertions or chromosomal translocations involving the IL-3 locus in any of our IL-3 expressing "mock" mutants from experiments LV3 and LV4 or the CNCG mutants C10, C40, C57 and C96.

Epigenetic changes could also be responsible for activation of IL-3 gene expression in these mutants. It was shown that site-specific demethylation of the IL-3 promoter is associated with IL-3 gene expression in activated mouse $CD8^+$ T cells (Fitzpatrick *et al.* 1998). We have not analysed the methylation status of the IL-3 promoter in our mutants, but demethylation could result in spontaneous IL-3 gene expression in these cells.

Lastly, it was shown that intracisternal A particle (IAP) transposition is a common mechanism of spontaneous transformation to IL-3 independence in the mouse IL-3 dependent D35 and FDC-P cell lines (Stocking *et al.* 1988; Heberlein *et al.* 1990; Wang *et al.* 1997). IAPs are transposable elements present in the genome of *Mus musculus* at 1000 copies per haploid genome (Lueders and Kuff 1980). They are expressed in embryonic and transformed mouse cells. IAPs are defective in envelope protein synthesis, and are retained in the endoplasmic reticulum of the cell (Perk and Dahlberg 1974). The IAP activity in BAF3 cells is not known, but could account for spontaneous transformation to IL-3 independence of these cells.

6.3.2. Phenotypical characterisation of the retroviral vector mutants

Mutant G18 that has a RIS near the IL-3 gene, and the spontaneous mutant C94, expressed the highest levels of IL-3 transcript. The lower levels at which the IL-3 transcript was expressed by the four IL-3 expressing CNCG mutants varied less than 10 fold.

We tested which of the 5 mutants from RV1 and RV2 were able to proliferate in serum free medium. Only mutants C96 and G18 were able to do so.

Conditioned medium from these 5 mutants was assayed for its ability to induce proliferation of parental BCL15 cells. An interesting finding was that only conditioned medium from mutant G18, but not C96, was able to support growth of parental BCL15 cells. The conditioned medium experiment suggests IL-3 is not secreted by mutant C96. It could be that this mutant proliferated independently of exogenous IL-3, because endogenously produced IL-3 binds the IL-3 receptor intracellularly. This has been reported for IL-3 that was engineered to be retained in the endoplasmic reticulum (Dunbar *et al.* 1989).

6.3.3. Why do retroviral vectors not transform BCL15 cells via Ghr activation?

Four fold more HV lentiviral (3.1 x 10^8 by Q-PCR) than retroviral (8.0 x 10^7 by Q-PCR) vector integrants were screened in the BCL15 cell line. Assuming random integration, this means there were fewer retroviral than lentiviral vector hits in the Ghr locus. This might be one explanation why no Ghr insertional mutants were obtained with retroviral vectors.

As was discussed at the end of chapter 4, the main mechanism whereby endogenous MLV (Akv) seemed to activate Ghr expression in a mouse model of pre-B cell leukaemia, was by enhancer insertion (Suzuki *et al.* 2002; Suzuki *et al.* 2006). The six RIS in the Ghr locus in these tumours clustered around the L2 promoter and two thirds were in reverse orientation to Ghr gene expression. The LTRs present in the vectors used in this study were derived from exogenous MLVs – Mo-MLV and MPSV. The enhancers contained within the LTRs of these vectors might not be able to trans-activate Ghr expression from the L2, L1 or L5 promoters. Another reason why Ghr enhancer activation was not observed, could be because vector integration in the BCL15 cell line did not occur near the L2, L1 and L5 transcription start sites.

The main mechanism whereby HV activates Ghr gene expression was shown to be by HIV-1 promoter insertion. Retroviral vectors integrated in the first intron of the Ghr in the same transcriptional orientation as Ghr gene transcription, should theoretically also be able to activate its expression by promoter insertion and generation of a spliced transcript. Even though MLV promoter insertion was described previously (Shen-Ong *et al.* 1986; Barker *et al.* 1992), it is expected to work less efficiently than HIV-1 promoter insertion, because the retroviral splice donor is less potent. Only ~30% of MLV primary transcripts are spliced to generate the env transcript. In comparison, in HIV-1 over 40 different mRNA transcripts are generated by alternative splicing from a single pre-mRNA transcript. All these transcripts are spliced from the major splice donor to multiple splice acceptors. Full length unspliced HIV-1 transcripts are only maintained once Rev is produced.

Chapter 7

7. Final discussion and future directions

The work described in this thesis was initiated following the news in 2002 that several recipients of gene therapy for X-SCID had developed leukaemia. Vector induced activation of known T cell proto-oncogenes was believed to be the driving force behind this process. In the light of these complications, the development of safer vectors for use in future gene therapy clinical trials was of the utmost importance. However, at the time no assay systems existed that allowed a quantitative assessment of the mutagenic potential of different vectors. Simple in vitro assays that would enable quick initial screening of such vectors were particularly desirable.

In this thesis I aimed to establish such an assay using the immortalised pro-B cell line BCL15, which is dependent on IL-3. The rate at which these cells became IL-3 independent following transduction with different vectors was measured. Below, I will discuss the general implications of the findings in this thesis. The assay will be compared to the Baum lab in vitro assay (Modlich *et al.* 2006; Zychlinski *et al.* 2008), discussed in section 1.5.2.3. of the introduction. Table 7.1. summarises the main differences between these two assays. Ways to improve our assay will be suggested. Lastly, adaptation of this assay into an in vivo screening tool for insertional mutagenesis will be considered.

Table 7.1. Comparison of the Collins versus Baum lab in vitro assays to measure insertional mutagenesis

	Collins Lab Assay			Baum Lab Assay		
Target cell	IL-3 dependent mouse pro-B lymphocyte cell line			Lin^- bone marrow from C57Bl6		
# of target cells	5×10^7			10^5		
MOI	4 – 12 (LV); 0.1 – 1 (RV)			20 (RV)		
Selectable Phenotype	IL-3 independence			Acquisition of serial replating ability		
Selection Protocol	2-3 day bulk culture with IL-3 Standard selection: • Wash of IL-3 on day 3/4 • Add back IL-3 on day 10/11 • Withdraw IL-3 day 13-14 • Score positive flasks Modified selection: • Wash of IL-3 on day 3/4 • Select in FCS plus bGH • Score positive wells			• 2 week bulk culture • Plate cells in 96 well plates at 100 cells/well • Score positive wells after 2 weeks		
Vectors tested & Integrant frequencies		Standard	Modified		Modlich	Zychlinski
	HV CSGW CNCG	4.3×10^{-8} 4.9×10^{-9} 9.5×10^{-8}	9.5×10^{-8} 3.3×10^{-9} $< 5.9 \times 10^{-8}$	WT SFFV SIN SFFV SIN SFFV cHS4 SIN EF1α	2.76×10^{-3} 2.30×10^{-4}	1.19×10^{-3} 4.70×10^{-4} 0
Target genes	Lenti: Ghr Retro: IL-3/ Bcl-X_L			Retro: Evi1		
Benefits	• No animal, recombinant cytokine cost • Assay scores independent clones • Sensitive assay for LV vector mutagenesis			• 1° cell is more relevant • Convenient readout • 10^4 fold more sensitive in detecting transforming events • Sensitive comparison between RV vector configurations		
Limitations	• Lower sensitivity • Inconvenient standard selection protocol • RV vector mutagenesis assay not very robust • Limited # of transforming target genes			• Long bulk culture phase • Assay does not score independent clones • Limited # of transforming target genes		

RV = retroviral; LV = lentiviral; # = number

7.1. Insights into the mechanisms of insertional mutagenesis by different vectors

At the outset of this project a number of target genes in the BAF3/BCL15 cell line had been identified in other studies whose activation could transform these cells to IL-3 independence. A gain of function assay for insertional mutagenesis should ideally screen as many loci as possible. The immortalisation of primary bone marrow cells by retroviral vectors in the Baum lab assay selects for Evi1 insertional mutants. In our assay, we similarly found targeting of specific genes by lentiviral versus retroviral vectors.

The main mechanism whereby the wild type HIV-1 LTR lentiviral vector HV transforms BCL15, is by activation of the Ghr. HV vector integration in these mutants invariably occurs in the first intron of this gene. A fusion transcript is present that originates from the HIV-1 5'LTR then splices from the HIV-1 major splice donor to the splice acceptor of Ghr exon 2. Ghr insertional mutants express the GHR on the cell surface and respond to GH in culture. This finding allowed the introduction of a modified, more sensitive selection protocol (See section 7.2.)

Activity of the HIV-1 5'LTR, which drives this mechanism of mutagenesis, is likely to be a phenomenon specific to the BAF3 cell line, which we plan to further investigate as discussed at the end of chapter 5. The HIV-1 LTR is unlikely to be active in the absence of Tat in most target cells of gene therapy. The clinical use of HIV-1 LTR-driven lentiviral vectors is likely to be restricted to gene therapy directed against HIV-1 infection. In the first clinical gene therapy trial using a lentiviral vector, the antisense against HIV-1 env was expressed in T cells in a Tat-dependent manner from the HIV-1 LTR (Levine *et al.* 2006). Use of such a vector in HIV-1 gene therapy to transduce more primitive long lived cells such as HSC might be more dangerous. In most gene therapy settings use of a SIN lentiviral vector will be far more desirable; especially as these can be produced transiently to similar titres as their wild-type counterparts. As demonstrated in this assay, the SIN configuration readily abolishes the promoter insertion mechanism of insertional mutagenesis. The residual transforming potential of SIN lentivectors

will largely depend on the choice of internal promoter, as demonstrated for SIN retroviral vectors in the Baum lab assay (Zychlinski *et al.* 2008). Unfortunately, characterisation of the CSGW mutants is incomplete, and the mechanism by which they gained IL-3 independence is unknown.

The mechanisms of transformation in IL-3 independent mutants obtained with retroviral vectors in this assay were not as clear-cut. In one mutant, a RIS near the IL-3 gene was found and activation of this gene presumably occurred via an enhancer effect. In the majority of retroviral vector mutants however, a role of the vector in transformation was not demonstrated, although insertional activation of genes that activate IL-3 expression remains a possibility.

7.2. Quantitation, Reproducibility and Applicability of the assay

Two experiments each were performed with the HV, CSGW and CNCG vectors. In some experiments mock mutants were obtained. The average frequency at which BCL15 cells spontaneously became IL-3 independent was 1.1×10^{-8}. IL-3 independent mutants were obtained at cell frequencies significantly over background with the HV and CNCG vectors. The average integrant frequencies at which mutants were obtained with the HV, CSGW and CNCG vectors were 4.3×10^{-8}, 4.9×10^{-9} and 9.5×10^{-8}, respectively. These frequencies were obtained from scoring independent clones. This is in contrast to the Baum lab assay; it was reported many of their clones were genetically identical and shared RIS.

Ideally, more repeat experiments for each vector would need to be performed to confirm these frequencies are reproducible and to determine variability in frequency between experiments. This was not possible, due to the labour-intensiveness of the standard selection protocol. A higher throughput selection protocol, see section 7.3., might make this feasible.

The current numbers allow us to conclude that the wild type HIV-1 LTR vector has a greater potential for insertional mutagenesis than the SIN vector CSGW; the numeric difference is 9 fold. Mechanistically, this is supported by the fact that none of the CSGW mutants activated Ghr expression. The difference in integrant frequencies between HV and CSGW was even greater (29 fold) after selecting

BCL15 cells transduced with these vectors in serum plus GH. This modified selection protocol is biased towards picking up Ghr-insertional mutants and is very sensitive in doing so; making it particularly suitable for screening safety optimised variants of HV vectors. The readout of this modified protocol is simple. A frequency can be calculated assuming that at least 70% of wells contain independent mutants. Another 1.9×10^8 HV integrants were recently screened by Abinav Gupta, a BSc student in the lab. 15 wells were scored. These clones have not been characterised yet, but IL-3 independent mutants were obtained at an integrant frequency of at least 5.5×10^{-8} in this experiment.

Four fold fewer retroviral than HV lentiviral integrants were screened. The average integrant frequency for CNCG was 9.5×10^{-8}, 2 fold higher than for HV. The CNCG integrant frequency is close to the 2×10^{-7} integrant frequency at which an MPSV retroviral vector transforms TF-1 cells to factor independence (Stocking *et al.* 1993). As discussed above, no one clear mechanism whereby retroviral vectors transform BCL15 cells to IL-3 independence was identified. So in contrast to HV mutagenesis, no tailor-made selection protocol that allowed rapid screening of different retroviral vector configurations could be introduced.

As shown in Table 7.1., the strength of the Baum lab assay is that it allows a direct comparison between the transforming potentials of different retroviral vector configurations (non-SIN versus SIN; different internal promoters within a SIN vector). Currently it is not possible to detect such differences in our cell line assay. Also going by the Baum lab assay numbers, we would have to screen 10 to 100 fold more retroviral integrants.

7.3. Towards a more high throughput selection protocol

At present only our modified selection protocol that is biased towards picking up Ghr-insertional mutants, has a convenient readout. This is much like the Baum lab assay where wells containing replating cells are scored 2 weeks after seeding. Our current standard selection protocol is time-consuming and limits the number of cells that can be screened in each experiment. This necessitates transduction at high MOI to screen a sufficient number of integrants. Resulting mutants contain several copies of the vector, increasing the number of vector integration sites that need to be cloned per mutant and complicating identification of the transforming gene. This was especially of practical importance in the early stages of the project, because we were interested in analysing the mechanisms of transformation. Ideally, more cells with on average one vector copy per cell could be screened. We need an unbiased selection protocol that allows mutants to grow out when grown in serum only without IL-3.

I started this thesis by testing assay conditions to facilitate selection for IL-3 independent mutants of the BAF3 cell line. X-irradiation was used as a strategy to eliminate more effectively parental BAF3 cells, so these would not be rescued upon re-addition of IL-3 to the media, which was needed to allow the expansion of IL-3 independent mutants from amidst dying cells. This strategy has not yet been tested on BCL15 cells. Bcl-2 expression should somewhat protect BCL15 cells from irradiation-induced apoptosis in the absence of IL-3 (Ascaso *et al.* 1994). Activation of cytokine or growth factor receptor expression should offer additive or synergistic protection to mutants. Currently, IL-3 independent mutants are selected in liquid medium. Selection cultures contain 5 x 10^5 cells, initially per well of a 24 well plate and later in a T25 flask. So far, I have never seen more than one clone growing out of a population of 5 x 10^5 cells. Selection in semi-solid media, such as soft agar or methylcellulose, would allow immediate selection for single cell clones. Outgrowth of mutants without the need for rescue with IL-3 might be possible in this setting, as this might not now be inhibited by the presence of large numbers of dead cells in the immediate vicinity. Similar to the current modified selection protocol, there would still need to be an initial

determination of the number of independent clones because the cells are grown as bulk culture for 3 days following transduction.

Another way to make the assay more sensitive is to introduce more biased selection, specifically for insertional mutants that have activated expression of receptors or subunits thereof. Selection of such mutants is at present limited by the availability of ligand of these receptors in serum. For example, selection in serum plus GM-CSF picked up retroviral insertional mutants of the FDC-P1(M) cell line that had activated expression of the GM-CSF receptor alpha chain (Laker *et al.* 2000). As demonstrated in the introduction to chapter 3, signalling by many other interleukin or growth factor receptors in BAF3 can substitute for IL-3.

7.4. Towards an in vivo assay

Neither parental BAF3 nor BCL15 cells are tumourigenic in vivo (Palacios and Steinmetz 1985). We have not established yet whether any of our IL-3 independent mutants are tumourigenic in vivo when injected into nude or syngeneic mice. In one study, wild-type MPL (WT-MPL) and a constitutively active mutant of MPL (Mu-MPL) were expressed in BAF3 cells (Alexander *et al.* 1995). WT-MPL cells were dependent on IL-3 or TPO for growth, whereas Mu-MPL cells grew independently of any growth factors. In another study, a different MPL mutant conferred serum-independent growth onto BAF3 cells (Onishi *et al.* 1996). In both studies WT-MPL and Mu-MPL expressing BAF3 cells were injected into BALB/c mice. Only mice that received Mu-MPL cells developed tumours after 2-3 weeks. Total follow-up of these animals was 2 – 3 months.

Like MPL, the GHR belongs to the haematopoietin family of receptors. BAF3 cells expressing GHR e.g. mutant HV A2, would therefore unlikely to be tumourigenic in vivo. This might be different for BCL15 cells expressing the GHR, as these cells have 2 oncogenic hits.

Adapting the in vitro mutagenesis assay in the BCL15 cell line to an in vivo model would be desirable as this is ultimately more relevant to the clinical gene therapy setting. BAF3 and/or BCL15 cells transduced with different vectors could be transplanted into nude or BALB/c mice that are then monitored for tumour

formation. Other factor-dependent cell lines could be used such as the IL-3-dependent cell lines FDCP-mix and M-07; or IL-2 dependent cell lines CTLL-2 and HT-2. These cell lines were not suitable for the development of an in vitro mutagenesis assay because they are not robust enough. It was demonstrated that nude and syngeneic mice injected with CTLL-2 cells transduced with a retroviral vector expressing IL-2, developed lymphomas (Yamada *et al.* 1987).

Lastly, to develop our in vitro assay we used an immortalised cell line expressing human Bcl-2. In extension, we plan to select for oncogenes cooperating with Bcl-2 by transducing factor-dependent cell lines or primary mouse bone marrow cells with vectors e.g. HV or MFG that encode human Bcl-2. These cells will then be transplanted back into mice and monitored for tumour formation. These type of studies have so far been performed using retroviral vectors encoding the Sox4 (Du *et al.* 2005) and SV40 large T antigen oncogenes (Li *et al.* 2007). In the former study, recipients of bone marrow transduced with the Sox4 expressing retroviral vector developed myeloid leukaemias (Du *et al.* 2005). In the latter study, both IL-3 dependent 32D as well as mouse bone marrow was transduced (Li *et al.* 2007). SV40 large T expression in 32D cells was insufficient to confer IL-3 independence on these cells. However, mice transplanted with these cells rapidly developed acute myeloid leukaemia (AML), whereas recipients of bone marrow transduced with the SV40 vector developed histiocytic sarcoma and myeloid leukaemias. As a proof of principle, in all tumours RIS in loci known to cooperate with SV40 large T antigen were identified.

References

Adam, M. A., N. Ramesh, A. D. Miller and W. R. Osborne (1991). "Internal initiation of translation in retroviral vectors carrying picornavirus 5' nontranslated regions." J Virol **65**(9): 4985-90.

Aiuti, A., B. Cassani, G. Andolfi, M. Mirolo, L. Biasco, A. Recchia, F. Urbinati, C. Valacca, S. Scaramuzza, M. Aker, S. Slavin, M. Cazzola, D. Sartori, A. Ambrosi, C. Di Serio, M. G. Roncarolo, F. Mavilio and C. Bordignon (2007). "Multilineage hematopoietic reconstitution without clonal selection in ADA-SCID patients treated with stem cell gene therapy." J Clin Invest **117**(8): 2233-40.

Aiuti, A., S. Slavin, M. Aker, F. Ficara, S. Deola, A. Mortellaro, S. Morecki, G. Andolfi, A. Tabucchi, F. Carlucci, E. Marinello, F. Cattaneo, S. Vai, P. Servida, R. Miniero, M. G. Roncarolo and C. Bordignon (2002). "Correction of ADA-SCID by stem cell gene therapy combined with nonmyeloablative conditioning." Science **296**(5577): 2410-3.

Akagi, K., T. Suzuki, R. M. Stephens, N. A. Jenkins and N. G. Copeland (2004). "RTCGD: retroviral tagged cancer gene database." Nucleic Acids Res **32**(Database issue): D523-7.

Alexander, W. S., D. Metcalf and A. R. Dunn (1995). "Point mutations within a dimer interface homology domain of c-Mpl induce constitutive receptor activity and tumorigenicity." Embo J **14**(22): 5569-78.

Ames, B. N., E. G. Gurney, J. A. Miller and H. Bartsch (1972). "Carcinogens as frameshift mutagens: metabolites and derivatives of 2-acetylaminofluorene and other aromatic amine carcinogens." Proc Natl Acad Sci U S A **69**(11): 3128-32.

Antoine, C., S. Muller, A. Cant, M. Cavazzana-Calvo, P. Veys, J. Vossen, A. Fasth, C. Heilmann, N. Wulffraat, R. Seger, S. Blanche, W. Friedrich, M. Abinun, G. Davies, R. Bredius, A. Schulz, P. Landais and A. Fischer (2003). "Long-term survival and transplantation of haemopoietic stem cells for immunodeficiencies: report of the European experience 1968-99." Lancet **361**(9357): 553-60.

Ascaso, R., J. Marvel, M. K. Collins and A. Lopez-Rivas (1994). "Interleukin-3 and Bcl-2 cooperatively inhibit etoposide-induced apoptosis in a murine pre-B cell line." Eur J Immunol **24**(3): 537-41.

Ashe, M. P., P. Griffin, W. James and N. J. Proudfoot (1995). "Poly(A) site selection in the HIV-1 provirus: inhibition of promoter-proximal

polyadenylation by the downstream major splice donor site." Genes Dev **9**(23): 3008-25.

Ashe, M. P., L. H. Pearson and N. J. Proudfoot (1997). "The HIV-1 5' LTR poly(A) site is inactivated by U1 snRNP interaction with the downstream major splice donor site." Embo J **16**(18): 5752-63.

Baixeras, E., S. Jeay, P. A. Kelly and M. C. Postel-Vinay (2001). "The proliferative and antiapoptotic actions of growth hormone and insulin-like growth factor-1 are mediated through distinct signaling pathways in the Pro-B Ba/F3 cell line." Endocrinology **142**(7): 2968-77.

Baker, S. J., S. G. Rane and E. P. Reddy (2007). "Hematopoietic cytokine receptor signaling." Oncogene **26**(47): 6724-37.

Baltimore, D. (1970). "RNA-dependent DNA polymerase in virions of RNA tumour viruses." Nature **226**(5252): 1209-11.

Barker, C. S., S. E. Bear, T. Keler, N. G. Copeland, D. J. Gilbert, N. A. Jenkins, R. S. Yeung and P. N. Tsichlis (1992). "Activation of the prolactin receptor gene by promoter insertion in a Moloney murine leukemia virus-induced rat thymoma." J Virol **66**(11): 6763-8.

Barre-Sinoussi, F., J. C. Chermann, F. Rey, M. T. Nugeyre, S. Chamaret, J. Gruest, C. Dauguet, C. Axler-Blin, F. Vezinet-Brun, C. Rouzioux, W. Rozenbaum and L. Montagnier (1983). "Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS)." Science **220**(4599): 868-71.

Bartholomew, C. and J. N. Ihle (1991). "Retroviral insertions 90 kilobases proximal to the Evi-1 myeloid transforming gene activate transcription from the normal promoter." Mol Cell Biol **11**(4): 1820-8.

Bartosch, B., J. Dubuisson and F. L. Cosset (2003). "Infectious hepatitis C virus pseudo-particles containing functional E1-E2 envelope protein complexes." J Exp Med **197**(5): 633-42.

Battini, J. L., J. E. Rasko and A. D. Miller (1999). "A human cell-surface receptor for xenotropic and polytropic murine leukemia viruses: possible role in G protein-coupled signal transduction." Proc Natl Acad Sci U S A **96**(4): 1385-90.

Baum, C., J. Dullmann, Z. Li, B. Fehse, J. Meyer, D. A. Williams and C. von Kalle (2003). "Side effects of retroviral gene transfer into hematopoietic stem cells." Blood **101**(6): 2099-114.

Baum, C., C. von Kalle, F. J. Staal, Z. Li, B. Fehse, M. Schmidt, F. Weerkamp, S. Karlsson, G. Wagemaker and D. A. Williams (2004). "Chance or

necessity? Insertional mutagenesis in gene therapy and its consequences." Mol Ther **9**(1): 5-13.

Ben-David, Y., A. Lavigueur, G. Y. Cheong and A. Bernstein (1990). "Insertional inactivation of the p53 gene during friend leukemia: a new strategy for identifying tumor suppressor genes." New Biol **2**(11): 1015-23.

Berns, A., M. Breuer, S. Verbeek and M. van Lohuizen (1989). "Transgenic mice as a means to study synergism between oncogenes." Int J Cancer Suppl **4**: 22-5.

Besancon, F., A. Atfi, C. Gespach, Y. E. Cayre and M. F. Bourgeade (1998). "Evidence for a role of NF-kappaB in the survival of hematopoietic cells mediated by interleukin 3 and the oncogenic TEL/platelet-derived growth factor receptor beta fusion protein." Proc Natl Acad Sci U S A **95**(14): 8081-6.

Besnier, C., Y. Takeuchi and G. Towers (2002). "Restriction of lentivirus in monkeys." Proc Natl Acad Sci U S A **99**(18): 11920-5.

Bhatia, M., J. C. Wang, U. Kapp, D. Bonnet and J. E. Dick (1997). "Purification of primitive human hematopoietic cells capable of repopulating immune-deficient mice." Proc Natl Acad Sci U S A **94**(10): 5320-5.

Bijl, J., M. Sauvageau, A. Thompson and G. Sauvageau (2005). "High incidence of proviral integrations in the Hoxa locus in a new model of E2a-PBX1-induced B-cell leukemia." Genes Dev **19**(2): 224-33.

Blaese, R. M., K. W. Culver, A. D. Miller, C. S. Carter, T. Fleisher, M. Clerici, G. Shearer, L. Chang, Y. Chiang, P. Tolstoshev, J. J. Greenblatt, S. A. Rosenberg, H. Klein, M. Berger, C. A. Mullen, W. J. Ramsey, L. Muul, R. A. Morgan and W. F. Anderson (1995). "T lymphocyte-directed gene therapy for ADA- SCID: initial trial results after 4 years." Science **270**(5235): 475-80.

Blalock, W. L., C. Weinstein-Oppenheimer, F. Chang, P. E. Hoyle, X. Y. Wang, P. A. Algate, R. A. Franklin, S. M. Oberhaus, L. S. Steelman and J. A. McCubrey (1999). "Signal transduction, cell cycle regulatory, and anti-apoptotic pathways regulated by IL-3 in hematopoietic cells: possible sites for intervention with anti-neoplastic drugs." Leukemia **13**(8): 1109-66.

Bonzon, C. and H. Fan (1999). "Moloney murine leukemia virus-induced preleukemic thymic atrophy and enhanced thymocyte apoptosis correlate with disease pathogenicity." J Virol **73**(3): 2434-41.

Bordignon, C., L. D. Notarangelo, N. Nobili, G. Ferrari, G. Casorati, P. Panina, E. Mazzolari, D. Maggioni, C. Rossi, P. Servida, A. G. Ugazio and F. Mavilio

(1995). "Gene therapy in peripheral blood lymphocytes and bone marrow for ADA- immunodeficient patients." Science **270**(5235): 470-5.

Boshoff, C. and R. Weiss (2002). "AIDS-related malignancies." Nat Rev Cancer **2**(5): 373-82.

Brightman, B. K., A. Rein, D. J. Trepp and H. Fan (1991). "An enhancer variant of Moloney murine leukemia virus defective in leukemogenesis does not generate detectable mink cell focus-inducing virus in vivo." Proc Natl Acad Sci U S A **88**(6): 2264-8.

Brown, B. D., B. Gentner, A. Cantore, S. Colleoni, M. Amendola, A. Zingale, A. Baccarini, G. Lazzari, C. Galli and L. Naldini (2007). "Endogenous microRNA can be broadly exploited to regulate transgene expression according to tissue, lineage and differentiation state." Nat Biotechnol **25**(12): 1457-67.

Brunet, A., A. Bonni, M. J. Zigmond, M. Z. Lin, P. Juo, L. S. Hu, M. J. Anderson, K. C. Arden, J. Blenis and M. E. Greenberg (1999). "Akt promotes cell survival by phosphorylating and inhibiting a Forkhead transcription factor." Cell **96**(6): 857-68.

Buchschacher, G. L., Jr. (2001). "Introduction to retroviruses and retroviral vectors." Somat Cell Mol Genet **26**(1-6): 1-11.

Bueren, J. A., G. Guenechea, J. A. Casado, M. L. Lamana and J. C. Segovia (2003). "Genetic modification of hematopoietic stem cells: recent advances in the gene therapy of inherited diseases." Arch Med Res **34**(6): 589-99.

Bulanova, E., V. Budagian, Z. Orinska, H. Krause, R. Paus and S. Bulfone-Paus (2003). "Mast cells express novel functional IL-15 receptor alpha isoforms." J Immunol **170**(10): 5045-55.

Burns, J. C., T. Friedmann, W. Driever, M. Burrascano and J. K. Yee (1993). "Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells." Proc Natl Acad Sci U S A **90**(17): 8033-7.

Carteau, S., C. Hoffmann and F. Bushman (1998). "Chromosome structure and human immunodeficiency virus type 1 cDNA integration: centromeric alphoid repeats are a disfavored target." J Virol **72**(5): 4005-14.

Cavazzana-Calvo, M., S. Hacein-Bey, G. de Saint Basile, F. Gross, E. Yvon, P. Nusbaum, F. Selz, C. Hue, S. Certain, J. L. Casanova, P. Bousso, F. L. Deist and A. Fischer (2000). "Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease." Science **288**(5466): 669-72.

Ceci, J. D., C. P. Patriotis, C. Tsatsanis, A. M. Makris, R. Kovatch, D. A. Swing, N. A. Jenkins, P. N. Tsichlis and N. G. Copeland (1997). "Tpl-2 is an oncogenic kinase that is activated by carboxy-terminal truncation." Genes Dev **11**(6): 688-700.

Chatis, P. A., C. A. Holland, J. W. Hartley, W. P. Rowe and N. Hopkins (1983). "Role for the 3' end of the genome in determining disease specificity of Friend and Moloney murine leukemia viruses." Proc Natl Acad Sci U S A **80**(14): 4408-11.

Ciuffi, A. and F. D. Bushman (2006). "Retroviral DNA integration: HIV and the role of LEDGF/p75." Trends Genet.

Ciuffi, A., M. Llano, E. Poeschla, C. Hoffmann, J. Leipzig, P. Shinn, J. R. Ecker and F. Bushman (2005). "A role for LEDGF/p75 in targeting HIV DNA integration." Nat Med **11**(12): 1287-9.

Ciuffi, A., R. S. Mitchell, C. Hoffmann, J. Leipzig, P. Shinn, J. R. Ecker and F. D. Bushman (2006). "Integration site selection by HIV-based vectors in dividing and growth-arrested IMR-90 lung fibroblasts." Mol Ther **13**(2): 366-73.

Collins, M. K., J. Downward, A. Miyajima, K. Maruyama, K. Arai and R. C. Mulligan (1988). "Transfer of functional EGF receptors to an IL3-dependent cell line." J Cell Physiol **137**(2): 293-8.

Collins, M. K., P. Malde, A. Miyajima, K. Arai, K. A. Smith and R. C. Mulligan (1990). "Evidence that the level of the p55 component of the interleukin (IL) 2 receptor can control IL 2 responsiveness in a murine IL 3-dependent cell." Eur J Immunol **20**(3): 573-8.

Collins, M. K., J. Marvel, P. Malde and A. Lopez-Rivas (1992). "Interleukin 3 protects murine bone marrow cells from apoptosis induced by DNA damaging agents." J Exp Med **176**(4): 1043-51.

Collins, M. K., G. R. Perkins, G. Rodriguez-Tarduchy, M. A. Nieto and A. Lopez-Rivas (1994). "Growth factors as survival factors: regulation of apoptosis." Bioessays **16**(2): 133-8.

Corey, S. J. and S. M. Anderson (1999). "Src-related protein tyrosine kinases in hematopoiesis." Blood **93**(1): 1-14.

Corneliussen, B., A. Thornell, B. Hallberg and T. Grundstrom (1991). "Helix-loop-helix transcriptional activators bind to a sequence in glucocorticoid response elements of retrovirus enhancers." J Virol **65**(11): 6084-93.

Cosset, F. L., Y. Takeuchi, J. L. Battini, R. A. Weiss and M. K. Collins (1995). "High-titer packaging cells producing recombinant retroviruses resistant to human serum." J Virol **69**(12): 7430-6.

Cui, J. W., Y. J. Li, A. Sarkar, J. Brown, Y. H. Tan, M. Premyslova, C. Michaud, N. Iscove, G. J. Wang and Y. Ben-David (2007). "Retroviral insertional activation of the Fli-3 locus in erythroleukemias encoding a cluster of microRNAs that convert Epo-induced differentiation to proliferation." Blood **110**(7): 2631-40.

Cullen, B. R., P. T. Lomedico and G. Ju (1984). "Transcriptional interference in avian retroviruses--implications for the promoter insertion model of leukaemogenesis." Nature **307**(5948): 241-5.

Cuypers, H. T., G. Selten, W. Quint, M. Zijlstra, E. R. Maandag, W. Boelens, P. van Wezenbeek, C. Melief and A. Berns (1984). "Murine leukemia virus-induced T-cell lymphomagenesis: integration of proviruses in a distinct chromosomal region." Cell **37**(1): 141-50.

Cuypers, H. T., G. C. Selten, M. Zijlstra, R. E. de Goede, C. J. Melief and A. J. Berns (1986). "Tumor progression in murine leukemia virus-induced T-cell lymphomas: monitoring clonal selections with viral and cellular probes." J Virol **60**(1): 230-41.

Dalgleish, A. G., P. C. Beverley, P. R. Clapham, D. H. Crawford, M. F. Greaves and R. A. Weiss (1984). "The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus." Nature **312**(5996): 763-7.

Danos, O. and R. C. Mulligan (1988). "Safe and efficient generation of recombinant retroviruses with amphotropic and ecotropic host ranges." Proc Natl Acad Sci U S A **85**(17): 6460-4.

Das, A. T., A. Harwig, M. M. Vrolijk and B. Berkhout (2007). "The TAR hairpin of human immunodeficiency virus type 1 can be deleted when not required for Tat-mediated activation of transcription." J Virol **81**(14): 7742-8.

Dave, U. P., N. A. Jenkins and N. G. Copeland (2004). "Gene therapy insertional mutagenesis insights." Science **303**(5656): 333.

Davis, B. R., B. K. Brightman, K. G. Chandy and H. Fan (1987). "Characterization of a preleukemic state induced by Moloney murine leukemia virus: evidence for two infection events during leukemogenesis." Proc Natl Acad Sci U S A **84**(14): 4875-9.

de Noronha, C. M., M. P. Sherman, H. W. Lin, M. V. Cavrois, R. D. Moir, R. D. Goldman and W. C. Greene (2001). "Dynamic disruptions in nuclear envelope architecture and integrity induced by HIV-1 Vpr." Science **294**(5544): 1105-8.

Deichmann, A., S. Hacein-Bey-Abina, M. Schmidt, A. Garrigue, M. H. Brugman, J. Hu, H. Glimm, G. Gyapay, B. Prum, C. C. Fraser, N. Fischer, K. Schwarzwaelder, M. L. Siegler, D. de Ridder, K. Pike-Overzet, S. J. Howe, A. J. Thrasher, G. Wagemaker, U. Abel, F. J. Staal, E. Delabesse, J. L. Villeval, B. Aronow, C. Hue, C. Prinz, M. Wissler, C. Klanke, J. Weissenbach, I. Alexander, A. Fischer, C. von Kalle and M. Cavazzana-Calvo (2007). "Vector integration is nonrandom and clustered and influences the fate of lymphopoiesis in SCID-X1 gene therapy." J Clin Invest **117**(8): 2225-32.

del Peso, L., M. Gonzalez-Garcia, C. Page, R. Herrera and G. Nunez (1997). "Interleukin-3-induced phosphorylation of BAD through the protein kinase Akt." Science **278**(5338): 687-9.

Delenda, C. (2004). "Lentiviral vectors: optimization of packaging, transduction and gene expression." J Gene Med **6 Suppl 1**: S125-38.

Demaison, C., K. Parsley, G. Brouns, M. Scherr, K. Battmer, C. Kinnon, M. Grez and A. J. Thrasher (2002). "High-level transduction and gene expression in hematopoietic repopulating cells using a human immunodeficiency [correction of imunodeficiency] virus type 1-based lentiviral vector containing an internal spleen focus forming virus promoter." Hum Gene Ther **13**(7): 803-13.

Demoulin, J. B., C. Uyttenhove, E. Van Roost, B. DeLestre, D. Donckers, J. Van Snick and J. C. Renauld (1996). "A single tyrosine of the interleukin-9 (IL-9) receptor is required for STAT activation, antiapoptotic activity, and growth regulation by IL-9." Mol Cell Biol **16**(9): 4710-6.

Deneka, M., A. Pelchen-Matthews, R. Byland, E. Ruiz-Mateos and M. Marsh (2007). "In macrophages, HIV-1 assembles into an intracellular plasma membrane domain containing the tetraspanins CD81, CD9, and CD53." J Cell Biol **177**(2): 329-41.

Deng, H., R. Liu, W. Ellmeier, S. Choe, D. Unutmaz, M. Burkhart, P. Di Marzio, S. Marmon, R. E. Sutton, C. M. Hill, C. B. Davis, S. C. Peiper, T. J. Schall, D. R. Littman and N. R. Landau (1996). "Identification of a major co-receptor for primary isolates of HIV-1." Nature **381**(6584): 661-6.

Dinerstein, H., F. Lago, L. Goujon, F. Ferrag, N. Esposito, J. Finidori, P. A. Kelly and M. C. Postel-Vinay (1995). "The proline-rich region of the GH receptor is essential for JAK2 phosphorylation, activation of cell proliferation, and gene transcription." Mol Endocrinol **9**(12): 1701-7.

Donsante, A., D. G. Miller, Y. Li, C. Vogler, E. M. Brunt, D. W. Russell and M. S. Sands (2007). "AAV vector integration sites in mouse hepatocellular carcinoma." Science **317**(5837): 477.

Dragic, T., V. Litwin, G. P. Allaway, S. R. Martin, Y. Huang, K. A. Nagashima, C. Cayanan, P. J. Maddon, R. A. Koup, J. P. Moore and W. A. Paxton (1996). "HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5." Nature **381**(6584): 667-73.

Du, Y., N. A. Jenkins and N. G. Copeland (2005). "Insertional mutagenesis identifies genes that promote the immortalization of primary bone marrow progenitor cells." Blood **106**(12): 3932-9.

Du, Y., S. E. Spence, N. A. Jenkins and N. G. Copeland (2005). "Cooperating cancer-gene identification through oncogenic-retrovirus-induced insertional mutagenesis." Blood **106**(7): 2498-505.

Duesberg, P. H. and P. K. Vogt (1970). "Differences between the ribonucleic acids of transforming and nontransforming avian tumor viruses." Proc Natl Acad Sci U S A **67**(4): 1673-80.

Dull, T., R. Zufferey, M. Kelly, R. J. Mandel, M. Nguyen, D. Trono and L. Naldini (1998). "A third-generation lentivirus vector with a conditional packaging system." J Virol **72**(11): 8463-71.

Dumon, S., S. C. Santos, F. Debierre-Grockiego, V. Gouilleux-Gruart, L. Cocault, C. Boucheron, P. Mollat, S. Gisselbrecht and F. Gouilleux (1999). "IL-3 dependent regulation of Bcl-xL gene expression by STAT5 in a bone marrow derived cell line." Oncogene **18**(29): 4191-9.

Dunbar, C. E., T. M. Browder, J. S. Abrams and A. W. Nienhuis (1989). "COOH-terminal-modified interleukin-3 is retained intracellularly and stimulates autocrine growth." Science **245**(4925): 1493-6.

Dvorin, J. D., P. Bell, G. G. Maul, M. Yamashita, M. Emerman and M. H. Malim (2002). "Reassessment of the roles of integrase and the central DNA flap in human immunodeficiency virus type 1 nuclear import." J Virol **76**(23): 12087-96.

Edelstein, M. L., M. R. Abedi and J. Wixon (2007). "Gene therapy clinical trials worldwide to 2007--an update." J Gene Med **9**(10): 833-42.

Edens, A. and F. Talamantes (1998). "Alternative processing of growth hormone receptor transcripts." Endocr Rev **19**(5): 559-82.

Eglitis, M. A., P. Kantoff, E. Gilboa and W. F. Anderson (1985). "Gene expression in mice after high efficiency retroviral-mediated gene transfer." Science **230**(4732): 1395-8.

Ellerman and Bang (1908). "Experimentelle Leukamie bei Hunern." Centralbl. Bakteriol **46**: 595–609.

Engelman, A. and R. Craigie (1992). "Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro." J Virol **66**(11): 6361-9.

Engelman, A., K. Mizuuchi and R. Craigie (1991). "HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer." Cell **67**(6): 1211-21.

Fan, D. P. and B. M. Sefton (1978). "The entry into host cells of Sindbis virus, vesicular stomatitis virus and Sendai virus." Cell **15**(3): 985-92.

Fassati, A. and S. P. Goff (1999). "Characterization of intracellular reverse transcription complexes of Moloney murine leukemia virus." J Virol **73**(11): 8919-25.

Fassati, A. and S. P. Goff (2001). "Characterization of intracellular reverse transcription complexes of human immunodeficiency virus type 1." J Virol **75**(8): 3626-35.

Fassati, A., D. Gorlich, I. Harrison, L. Zaytseva and J. M. Mingot (2003). "Nuclear import of HIV-1 intracellular reverse transcription complexes is mediated by importin 7." Embo J **22**(14): 3675-85.

Feng, Y., C. C. Broder, P. E. Kennedy and E. A. Berger (1996). "HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor." Science **272**(5263): 872-7.

Fitzpatrick, D. R., K. M. Shirley, L. E. McDonald, H. Bielefeldt-Ohmann, G. F. Kay and A. Kelso (1998). "Distinct methylation of the interferon gamma (IFN-gamma) and interleukin 3 (IL-3) genes in newly activated primary CD8+ T lymphocytes: regional IFN-gamma promoter demethylation and mRNA expression are heritable in CD44(high)CD8+ T cells." J Exp Med **188**(1): 103-17.

Follenzi, A., M. Battaglia, A. Lombardo, A. Annoni, M. G. Roncarolo and L. Naldini (2004). "Targeting lentiviral vector expression to hepatocytes limits transgene-specific immune response and establishes long-term expression of human antihemophilic factor IX in mice." Blood **103**(10): 3700-9.

Fred C. Krebs, T. H. H., Shane Quiterio, Suzanne Gartner, and Brian Wigdahl (2001). "Lentiviral LTR-directed Expression, Sequence Variation, and Disease Pathogenesis." Human Retroviruses and AIDS 2001 Compendium http://www.hiv.lanl.gov/content/sequence/HIV/COMPENDIUM/2001/partI/Wigdahl.pdf.

Freed, E. O. and A. J. Mouland (2006). "The cell biology of HIV-1 and other retroviruses." Retrovirology **3**: 77.

Furger, A., J. Monks and N. J. Proudfoot (2001). "The retroviruses human immunodeficiency virus type 1 and Moloney murine leukemia virus adopt radically different strategies to regulate promoter-proximal polyadenylation." J Virol **75**(23): 11735-46.

Garcia-Martinez, L. F., G. Mavankal, J. M. Neveu, W. S. Lane, D. Ivanov and R. B. Gaynor (1997). "Purification of a Tat-associated kinase reveals a TFIIH complex that modulates HIV-1 transcription." Embo J **16**(10): 2836-50.

Gaspar, H. B., E. Bjorkegren, K. Parsley, K. C. Gilmour, D. King, J. Sinclair, F. Zhang, A. Giannakopoulos, S. Adams, L. D. Fairbanks, J. Gaspar, L. Henderson, J. H. Xu-Bayford, E. G. Davies, P. A. Veys, C. Kinnon and A. J. Thrasher (2006). "Successful reconstitution of immunity in ADA-SCID by stem cell gene therapy following cessation of PEG-ADA and use of mild preconditioning." Mol Ther **14**(4): 505-13.

Gaspar, H. B., K. L. Parsley, S. Howe, D. King, K. C. Gilmour, J. Sinclair, G. Brouns, M. Schmidt, C. Von Kalle, T. Barington, M. A. Jakobsen, H. O. Christensen, A. Al Ghonaium, H. N. White, J. L. Smith, R. J. Levinsky, R. R. Ali, C. Kinnon and A. J. Thrasher (2004). "Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector." Lancet **364**(9452): 2181-7.

Gatti, R. A., H. J. Meuwissen, H. D. Allen, R. Hong and R. A. Good (1968). "Immunological reconstitution of sex-linked lymphopenic immunological deficiency." Lancet **2**(7583): 1366-9.

Gillis, P. and J. S. Malter (1991). "The adenosine-uridine binding factor recognizes the AU-rich elements of cytokine, lymphokine, and oncogene mRNAs." J Biol Chem **266**(5): 3172-7.

Golemis, E. A., N. A. Speck and N. Hopkins (1990). "Alignment of U3 region sequences of mammalian type C viruses: identification of highly conserved motifs and implications for enhancer design." J Virol **64**(2): 534-42.

Gottlinger, H. G. (2001). "The HIV-1 assembly machine." Aids **15 Suppl 5**: S13-20.

Gouilleux, F., C. Pallard, I. Dusanter-Fourt, H. Wakao, L. A. Haldosen, G. Norstedt, D. Levy and B. Groner (1995). "Prolactin, growth hormone, erythropoietin and granulocyte-macrophage colony stimulating factor induce MGF-Stat5 DNA binding activity." Embo J **14**(9): 2005-13.

Greenberger, J. S., R. J. Eckner, W. Ostertag, G. Colletta, S. Boschetti, H. Nagasawa, A. Karpas, R. R. Weichselbaum and W. C. Moloney (1980).

"Release of spleen focus-forming virus (SFFV) from differentiation inducible promyelocytic leukemia cell lines transformed in vitro by Friend leukemia virus." Virology **105**(2): 425-35.

Grosovsky, A. J., A. Skandalis, L. Hasegawa and B. N. Walter (1993). "Insertional inactivation of the tk locus in a human B lymphoblastoid cell line by a retroviral shuttle vector." Mutat Res **289**(2): 297-308.

Grossman, Z., M. Meier-Schellersheim, W. E. Paul and L. J. Picker (2006). "Pathogenesis of HIV infection: what the virus spares is as important as what it destroys." Nat Med **12**(3): 289-95.

Gunther, C. V. and B. J. Graves (1994). "Identification of ETS domain proteins in murine T lymphocytes that interact with the Moloney murine leukemia virus enhancer." Mol Cell Biol **14**(11): 7569-80.

Gutierrez del Arroyo, A., C. Gil-Lamagniere, I. Lazaro, M. C. de Marco, I. Layunta and A. Silva (2000). "Involvement of p53 and interleukin 3 in the up-regulation of CD95 (APO-1/Fas) by X-ray irradiation." Oncogene **19**(32): 3647-55.

Haase, A. T. (2005). "Perils at mucosal front lines for HIV and SIV and their hosts." Nat Rev Immunol **5**(10): 783-92.

Hacein-Bey-Abina, S., F. Le Deist, F. Carlier, C. Bouneaud, C. Hue, J. P. De Villartay, A. J. Thrasher, N. Wulffraat, R. Sorensen, S. Dupuis-Girod, A. Fischer, E. G. Davies, W. Kuis, L. Leiva and M. Cavazzana-Calvo (2002). "Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy." N Engl J Med **346**(16): 1185-93.

Hacein-Bey-Abina, S., C. Von Kalle, M. Schmidt, M. P. McCormack, N. Wulffraat, P. Leboulch, A. Lim, C. S. Osborne, R. Pawliuk, E. Morillon, R. Sorensen, A. Forster, P. Fraser, J. I. Cohen, G. de Saint Basile, I. Alexander, U. Wintergerst, T. Frebourg, A. Aurias, D. Stoppa-Lyonnet, S. Romana, I. Radford-Weiss, F. Gross, F. Valensi, E. Delabesse, E. Macintyre, F. Sigaux, J. Soulier, L. E. Leiva, M. Wissler, C. Prinz, T. H. Rabbitts, F. Le Deist, A. Fischer and M. Cavazzana-Calvo (2003). "LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1." Science **302**(5644): 415-9.

Hahn, W. C. and R. A. Weinberg (2002). "Modelling the molecular circuitry of cancer." Nat Rev Cancer **2**(5): 331-41.

Hanania, E. G., R. E. Giles, J. Kavanagh, S. Q. Fu, D. Ellerson, Z. Zu, T. Wang, Y. Su, A. Kudelka, Z. Rahman, F. Holmes, G. Hortobagyi, D. Claxton, C. Bachier, P. Thall, S. Cheng, J. Hester, J. M. Ostrove, R. E. Bird, A. Chang, M. Korbling, D. Seong, R. Cote, T. Holzmayer, A. B. Deisseroth and et al. (1996). "Results of MDR-1 vector modification trial indicate that

granulocyte/macrophage colony-forming unit cells do not contribute to posttransplant hematopoietic recovery following intensive systemic therapy." Proc Natl Acad Sci U S A **93**(26): 15346-51.

Hannemann, J., T. Hara, M. Kawai, A. Miyajima, W. Ostertag and C. Stocking (1995). "Sequential mutations in the interleukin-3 (IL3)/granulocyte-macrophage colony-stimulating factor/IL5 receptor beta-subunit genes are necessary for the complete conversion to growth autonomy mediated by a truncated beta C subunit." Mol Cell Biol **15**(5): 2402-12.

Harada, F., G. G. Peters and J. E. Dahlberg (1979). "The primer tRNA for Moloney murine leukemia virus DNA synthesis. Nucleotide sequence and aminoacylation of tRNAPro." J Biol Chem **254**(21): 10979-85.

Hartley, J. W., N. K. Wolford, L. J. Old and W. P. Rowe (1977). "A new class of murine leukemia virus associated with development of spontaneous lymphomas." Proc Natl Acad Sci U S A **74**(2): 789-92.

Hatziioannou, T., E. Delahaye, F. Martin, S. J. Russell and F. L. Cosset (1999). "Retroviral display of functional binding domains fused to the amino terminus of influenza hemagglutinin." Hum Gene Ther **10**(9): 1533-44.

Hatziioannou, T., S. Valsesia-Wittmann, S. J. Russell and F. L. Cosset (1998). "Incorporation of fowl plague virus hemagglutinin into murine leukemia virus particles and analysis of the infectivity of the pseudotyped retroviruses." J Virol **72**(6): 5313-7.

Hayward, W. S., B. G. Neel and S. M. Astrin (1981). "Activation of a cellular onc gene by promoter insertion in ALV-induced lymphoid leukosis." Nature **290**(5806): 475-80.

Heberlein, C., M. Kawai, M. J. Franz, G. Beck-Engeser, C. P. Daniel, W. Ostertag and C. Stocking (1990). "Retrotransposons as mutagens in the induction of growth autonomy in hematopoietic cells." Oncogene **5**(12): 1799-807.

Heeney, J. L., A. G. Dalgleish and R. A. Weiss (2006). "Origins of HIV and the evolution of resistance to AIDS." Science **313**(5786): 462-6.

Hematti, P., B. K. Hong, C. Ferguson, R. Adler, H. Hanawa, S. Sellers, I. E. Holt, C. E. Eckfeldt, Y. Sharma, M. Schmidt, C. von Kalle, D. A. Persons, E. M. Billings, C. M. Verfaillie, A. W. Nienhuis, T. G. Wolfsberg, C. E. Dunbar and B. Calmels (2004). "Distinct genomic integration of MLV and SIV vectors in primate hematopoietic stem and progenitor cells." PLoS Biol **2**(12): e423.

Herndier, B. G., B. T. Shiramizu, N. E. Jewett, K. D. Aldape, G. R. Reyes and M. S. McGrath (1992). "Acquired immunodeficiency syndrome-associated T-

cell lymphoma: evidence for human immunodeficiency virus type 1-associated T-cell transformation." Blood **79**(7): 1768-74.

Higashimoto, T., F. Urbinati, A. Perumbeti, G. Jiang, A. Zarzuela, L. J. Chang, D. B. Kohn and P. Malik (2007). "The woodchuck hepatitis virus post-transcriptional regulatory element reduces readthrough transcription from retroviral vectors." Gene Ther **14**(17): 1298-304.

Himmel, K. L., F. Bi, H. Shen, N. A. Jenkins, N. G. Copeland, Y. Zheng and D. A. Largaespada (2002). "Activation of clg, a novel dbl family guanine nucleotide exchange factor gene, by proviral insertion at evi24, a common integration site in B cell and myeloid leukemias." J Biol Chem **277**(16): 13463-72.

Hindmarsh, P. and J. Leis (1999). "Retroviral DNA integration." Microbiol Mol Biol Rev **63**(4): 836-43, table of contents.

Holman, A. G. and J. M. Coffin (2005). "Symmetrical base preferences surrounding HIV-1, avian sarcoma/leukosis virus, and murine leukemia virus integration sites." Proc Natl Acad Sci U S A **102**(17): 6103-7.

Holmes, R. K., M. H. Malim and K. N. Bishop (2007). "APOBEC-mediated viral restriction: not simply editing?" Trends Biochem Sci **32**(3): 118-28.

Hoover, R. R., M. J. Gerlach, E. Y. Koh and G. Q. Daley (2001). "Cooperative and redundant effects of STAT5 and Ras signaling in BCR/ABL transformed hematopoietic cells." Oncogene **20**(41): 5826-35.

Horn, S., J. Meyer, C. Stocking, W. Ostertag and M. Jucker (2003). "An increase in the expression and total activity of endogenous p60(c-Src) in several factor-independent mutants of a human GM-CSF-dependent leukemia cell line (TF-1)." Oncogene **22**(46): 7170-80.

Ikeda, Y., Y. Takeuchi, F. Martin, F. L. Cosset, K. Mitrophanous and M. Collins (2003). "Continuous high-titer HIV-1 vector production." Nat Biotechnol **21**(5): 569-72.

Irusta, P. M. and D. DiMaio (1998). "A single amino acid substitution in a WW-like domain of diverse members of the PDGF receptor subfamily of tyrosine kinases causes constitutive receptor activation." Embo J **17**(23): 6912-23.

Jacque, J. M. and M. Stevenson (2006). "The inner-nuclear-envelope protein emerin regulates HIV-1 infectivity." Nature **441**(7093): 641-5.

Jeay, S., G. E. Sonenshein, P. A. Kelly, M. C. Postel-Vinay and E. Baixeras (2001). "Growth hormone exerts antiapoptotic and proliferative effects through two different pathways involving nuclear factor-kappaB and phosphatidylinositol 3-kinase." Endocrinology **142**(1): 147-56.

Jeay, S., G. E. Sonenshein, M. C. Postel-Vinay and E. Baixeras (2000). "Growth hormone prevents apoptosis through activation of nuclear factor-kappaB in interleukin-3-dependent Ba/F3 cell line." Mol Endocrinol **14**(5): 650-61.

Jiang, M., J. Mak, A. Ladha, E. Cohen, M. Klein, B. Rovinski and L. Kleiman (1993). "Identification of tRNAs incorporated into wild-type and mutant human immunodeficiency virus type 1." J Virol **67**(6): 3246-53.

Johnson, P. and S. Benchimol (1992). "Friend virus induced murine erythroleukaemia: the p53 locus." Cancer Surv **12**: 137-51.

Jones, K. A., J. T. Kadonaga, P. A. Luciw and R. Tjian (1986). "Activation of the AIDS retrovirus promoter by the cellular transcription factor, Sp1." Science **232**(4751): 755-9.

Jones, S. S., A. D. D'Andrea, L. L. Haines and G. G. Wong (1990). "Human erythropoietin receptor: cloning, expression, and biologic characterization." Blood **76**(1): 31-5.

Jonkers, J. and A. Berns (1996). "Retroviral insertional mutagenesis as a strategy to identify cancer genes." Biochim Biophys Acta **1287**(1): 29-57.

Kao, S. Y., A. F. Calman, P. A. Luciw and B. M. Peterlin (1987). "Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product." Nature **330**(6147): 489-93.

Kasprzycka, M., M. Majewski, Z. J. Wang, A. Ptasznik, M. Wysocka, Q. Zhang, M. Marzec, P. Gimotty, M. R. Crompton and M. A. Wasik (2006). "Expression and oncogenic role of Brk (PTK6/Sik) protein tyrosine kinase in lymphocytes." Am J Pathol **168**(5): 1631-41.

Katz, R. A. and A. M. Skalka (1994). "The retroviral enzymes." Annu Rev Biochem **63**: 133-73.

Killebrew, D. and B. Shiramizu (2004). "Pathogenesis of HIV-associated non-Hodgkin lymphoma." Curr HIV Res **2**(3): 215-21.

Kim, J. W., E. I. Closs, L. M. Albritton and J. M. Cunningham (1991). "Transport of cationic amino acids by the mouse ecotropic retrovirus receptor." Nature **352**(6337): 725-8.

Kim, Y. K., C. F. Bourgeois, R. Pearson, M. Tyagi, M. J. West, J. Wong, S. Y. Wu, C. M. Chiang and J. Karn (2006). "Recruitment of TFIIH to the HIV LTR is a rate-limiting step in the emergence of HIV from latency." Embo J **25**(15): 3596-604.

King, W., M. D. Patel, L. I. Lobel, S. P. Goff and M. C. Nguyen-Huu (1985). "Insertion mutagenesis of embryonal carcinoma cells by retroviruses." Science **228**(4699): 554-8.

Kinoshita, S., L. Su, M. Amano, L. A. Timmerman, H. Kaneshima and G. P. Nolan (1997). "The T cell activation factor NF-ATc positively regulates HIV-1 replication and gene expression in T cells." Immunity **6**(3): 235-44.

Klatzmann, D., E. Champagne, S. Chamaret, J. Gruest, D. Guetard, T. Hercend, J. C. Gluckman and L. Montagnier (1984). "T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV." Nature **312**(5996): 767-8.

Ko, C. B., B. S. Lee, S. H. Cha, D. Sul, S. G. Paik and H. S. Kang (2007). "Distinct role of IL-3 promoter and enhancer region in murine mast cells." Mol Immunol **44**(7): 1569-76.

Kohn, D. B., K. I. Weinberg, J. A. Nolta, L. N. Heiss, C. Lenarsky, G. M. Crooks, M. E. Hanley, G. Annett, J. S. Brooks, A. el-Khoureiy and et al. (1995). "Engraftment of gene-modified umbilical cord blood cells in neonates with adenosine deaminase deficiency." Nat Med **1**(10): 1017-23.

Krall, W. J., D. C. Skelton, X. J. Yu, I. Riviere, P. Lehn, R. C. Mulligan and D. B. Kohn (1996). "Increased levels of spliced RNA account for augmented expression from the MFG retroviral vector in hematopoietic cells." Gene Ther **3**(1): 37-48.

Kustikova, O., B. Fehse, U. Modlich, M. Yang, J. Dullmann, K. Kamino, N. von Neuhoff, B. Schlegelberger, Z. Li and C. Baum (2005). "Clonal dominance of hematopoietic stem cells triggered by retroviral gene marking." Science **308**(5725): 1171-4.

Laker, C., J. Friel, M. J. Franz, T. Hara, P. Papadopoulos, W. Ostertag and C. Stocking (2000). "Transcriptional activation of the granulocyte-macrophage colony-stimulating factor receptor gene in cell mutants." Exp Cell Res **259**(1): 1-11.

Lazo, P. A., J. S. Lee and P. N. Tsichlis (1990). "Long-distance activation of the Myc protooncogene by provirus insertion in Mlvi-1 or Mlvi-4 in rat T-cell lymphomas." Proc Natl Acad Sci U S A **87**(1): 170-3.

Leverrier, Y., J. Thomas, G. R. Perkins, M. Mangeney, M. K. Collins and J. Marvel (1997). "In bone marrow derived Baf-3 cells, inhibition of apoptosis by IL-3 is mediated by two independent pathways." Oncogene **14**(4): 425-30.

Levine, B. L., L. M. Humeau, J. Boyer, R. R. MacGregor, T. Rebello, X. Lu, G. K. Binder, V. Slepushkin, F. Lemiale, J. R. Mascola, F. D. Bushman, B.

Dropulic and C. H. June (2006). "Gene transfer in humans using a conditionally replicating lentiviral vector." Proc Natl Acad Sci U S A **103**(46): 17372-7.

Lewinski, M. K., M. Yamashita, M. Emerman, A. Ciuffi, H. Marshall, G. Crawford, F. Collins, P. Shinn, J. Leipzig, S. Hannenhalli, C. C. Berry, J. R. Ecker and F. D. Bushman (2006). "Retroviral DNA integration: viral and cellular determinants of target-site selection." PLoS Pathog **2**(6): e60.

Lewis, A. F., T. Stacy, W. R. Green, L. Taddesse-Heath, J. W. Hartley and N. A. Speck (1999). "Core-binding factor influences the disease specificity of Moloney murine leukemia virus." J Virol **73**(7): 5535-47.

Lewis, P., M. Hensel and M. Emerman (1992). "Human immunodeficiency virus infection of cells arrested in the cell cycle." Embo J **11**(8): 3053-8.

Li, J. and A. C. Sartorelli (1995). "Granulocyte colony-stimulating factor receptor mediated cell growth and differentiation." Oncol Res **7**(12): 591-601.

Li, Q., L. Duan, J. D. Estes, Z. M. Ma, T. Rourke, Y. Wang, C. Reilly, J. Carlis, C. J. Miller and A. T. Haase (2005). "Peak SIV replication in resting memory CD4+ T cells depletes gut lamina propria CD4+ T cells." Nature **434**(7037): 1148-52.

Li, Q. X. and H. Fan (1990). "Combined infection by Moloney murine leukemia virus and a mink cell focus-forming virus recombinant induces cytopathic effects in fibroblasts or in long-term bone marrow cultures from preleukemic mice." J Virol **64**(8): 3701-11.

Li, Y., E. Golemis, J. W. Hartley and N. Hopkins (1987). "Disease specificity of nondefective Friend and Moloney murine leukemia viruses is controlled by a small number of nucleotides." J Virol **61**(3): 693-700.

Li, Z., J. Dullmann, B. Schiedlmeier, M. Schmidt, C. von Kalle, J. Meyer, M. Forster, C. Stocking, A. Wahlers, O. Frank, W. Ostertag, K. Kuhlcke, H. G. Eckert, B. Fehse and C. Baum (2002). "Murine leukemia induced by retroviral gene marking." Science **296**(5567): 497.

Li, Z., O. S. Kustikova, K. Kamino, T. Neumann, M. Rhein, E. Grassman, B. Fehse and C. Baum (2007). "Insertional mutagenesis by replication-deficient retroviral vectors encoding the large T oncogene." Ann N Y Acad Sci **1106**: 95-113.

Logan, A. C., D. L. Haas, T. Kafri and D. B. Kohn (2004). "Integrated self-inactivating lentiviral vectors produce full-length genomic transcripts competent for encapsidation and integration." J Virol **78**(16): 8421-36.

Low, W., G. Olmos-Centenera, C. Madsen, Y. Leverrier and M. K. Collins (2001). "Role of Bax in apoptosis of IL-3-dependent cells." Oncogene **20**(33): 4476-83.

Lueders, K. K. and E. L. Kuff (1980). "Intracisternal A-particle genes: identification in the genome of Mus musculus and comparison of multiple isolates from a mouse gene library." Proc Natl Acad Sci U S A **77**(6): 3571-5.

Lund, A. H., G. Turner, A. Trubetskoy, E. Verhoeven, E. Wientjens, D. Hulsman, R. Russell, R. A. DePinho, J. Lenz and M. van Lohuizen (2002). "Genome-wide retroviral insertional tagging of genes involved in cancer in Cdkn2a-deficient mice." Nat Genet **32**(1): 160-5.

Malde, P. and M. K. Collins (1994). "Disregulation of Myc expression in murine bone marrow cells results in an inability to proliferate in sub-optimal growth factor and an increased sensitivity to DNA damage." Int Immunol **6**(8): 1169-76.

Malim, M. H., J. Hauber, S. Y. Le, J. V. Maizel and B. R. Cullen (1989). "The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA." Nature **338**(6212): 254-7.

Malinge, S., R. Ben-Abdelali, C. Settegrana, I. Radford-Weiss, M. Debre, K. Beldjord, E. A. Macintyre, J. L. Villeval, W. Vainchenker, R. Berger, O. A. Bernard, E. Delabesse and V. Penard-Lacronique (2007). "Novel activating JAK2 mutation in a patient with Down syndrome and B-cell precursor acute lymphoblastic leukemia." Blood **109**(5): 2202-4.

Mancebo, H. S., G. Lee, J. Flygare, J. Tomassini, P. Luu, Y. Zhu, J. Peng, C. Blau, D. Hazuda, D. Price and O. Flores (1997). "P-TEFb kinase is required for HIV Tat transcriptional activation in vivo and in vitro." Genes Dev **11**(20): 2633-44.

Manley, N. R., M. O'Connell, W. Sun, N. A. Speck and N. Hopkins (1993). "Two factors that bind to highly conserved sequences in mammalian type C retroviral enhancers." J Virol **67**(4): 1967-75.

Mann, R., R. C. Mulligan and D. Baltimore (1983). "Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus." Cell **33**(1): 153-9.

Marin, M., C. S. Tailor, A. Nouri, S. L. Kozak and D. Kabat (1999). "Polymorphisms of the cell surface receptor control mouse susceptibilities to xenotropic and polytropic leukemia viruses." J Virol **73**(11): 9362-8.

Markowitz, D., S. Goff and A. Bank (1988). "Construction and use of a safe and efficient amphotropic packaging cell line." Virology **167**(2): 400-6.

Martin-Serrano, J. (2007). "The role of ubiquitin in retroviral egress." Traffic **8**(10): 1297-303.

Martin, G. S. (1970). "Rous sarcoma virus: a function required for the maintenance of the transformed state." Nature **227**(5262): 1021-3.

Martinez-Moczygemba, M. and D. P. Huston (2003). "Biology of common beta receptor-signaling cytokines: IL-3, IL-5, and GM-CSF." J Allergy Clin Immunol **112**(4): 653-65; quiz 66.

Mathey-Prevot, B., G. Nabel, R. Palacios and D. Baltimore (1986). "Abelson virus abrogation of interleukin-3 dependence in a lymphoid cell line." Mol Cell Biol **6**(11): 4133-5.

Mathieu, A. L., S. Gonin, Y. Leverrier, B. Blanquier, J. Thomas, C. Dantin, G. Martin, G. Baverel and J. Marvel (2001). "Activation of the phosphatidylinositol 3-kinase/Akt pathway protects against interleukin-3 starvation but not DNA damage-induced apoptosis." J Biol Chem **276**(14): 10935-42.

Mattapallil, J. J., D. C. Douek, B. Hill, Y. Nishimura, M. Martin and M. Roederer (2005). "Massive infection and loss of memory CD4+ T cells in multiple tissues during acute SIV infection." Nature **434**(7037): 1093-7.

Mayo, M. W., X. Y. Wang, P. A. Algate, G. F. Arana, P. E. Hoyle, L. S. Steelman and J. A. McCubrey (1995). "Synergy between AUUUA motif disruption and enhancer insertion results in autocrine transformation of interleukin-3-dependent hematopoietic cells." Blood **86**(8): 3139-50.

McClure, M. O., M. Marsh and R. A. Weiss (1988). "Human immunodeficiency virus infection of CD4-bearing cells occurs by a pH-independent mechanism." Embo J **7**(2): 513-8.

McClure, M. O., M. A. Sommerfelt, M. Marsh and R. A. Weiss (1990). "The pH independence of mammalian retrovirus infection." J Gen Virol **71 (Pt 4)**: 767-73.

McDonald, D., L. Wu, S. M. Bohks, V. N. KewalRamani, D. Unutmaz and T. J. Hope (2003). "Recruitment of HIV and its receptors to dendritic cell-T cell junctions." Science **300**(5623): 1295-7.

Menon, R. K., A. Shaufl, J. H. Yu, D. A. Stephan and R. P. Friday (2001). "Identification and characterization of a novel transcript of the murine growth hormone receptor gene exhibiting development- and tissue-specific expression." Mol Cell Endocrinol **172**(1-2): 135-46.

Menon, R. K., D. A. Stephan, M. Singh, S. M. Morris, Jr. and L. Zou (1995). "Cloning of the promoter-regulatory region of the murine growth hormone

receptor gene. Identification of a developmentally regulated enhancer element." J Biol Chem **270**(15): 8851-9.

Meyer, J., C. Laker, N. Janzir, M. J. Franz, U. Bergholz, W. Ostertag and C. Stocking (2002). "Activation of the gene for the PDGF receptor beta1 (PDGFRbeta) in interleukin-3-dependent myeloid cells by retroviral insertional mutagenesis: implications for the transforming potential of PDGFRbeta." Growth Factors **20**(3): 131-40.

Miller, A. D. and C. Buttimore (1986). "Redesign of retrovirus packaging cell lines to avoid recombination leading to helper virus production." Mol Cell Biol **6**(8): 2895-902.

Miller, A. D., J. V. Garcia, N. von Suhr, C. M. Lynch, C. Wilson and M. V. Eiden (1991). "Construction and properties of retrovirus packaging cells based on gibbon ape leukemia virus." J Virol **65**(5): 2220-4.

Mitchell, R. S., B. F. Beitzel, A. R. Schroder, P. Shinn, H. Chen, C. C. Berry, J. R. Ecker and F. D. Bushman (2004). "Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences." PLoS Biol **2**(8): E234.

Miyazaki, T., Z. J. Liu, A. Kawahara, Y. Minami, K. Yamada, Y. Tsujimoto, E. L. Barsoumian, R. M. Permutter and T. Taniguchi (1995). "Three distinct IL-2 signaling pathways mediated by bcl-2, c-myc, and lck cooperate in hematopoietic cell proliferation." Cell **81**(2): 223-31.

Modlich, U., J. Bohne, M. Schmidt, C. von Kalle, S. Knoss, A. Schambach and C. Baum (2006). "Cell-culture assays reveal the importance of retroviral vector design for insertional genotoxicity." Blood **108**(8): 2545-53.

Modlich, U., O. S. Kustikova, M. Schmidt, C. Rudolph, J. Meyer, Z. Li, K. Kamino, N. von Neuhoff, B. Schlegelberger, K. Kuehlcke, K. D. Bunting, S. Schmidt, A. Deichmann, C. von Kalle, B. Fehse and C. Baum (2005). "Leukemias following retroviral transfer of multidrug resistance 1 (MDR1) are driven by combinatorial insertional mutagenesis." Blood **105**(11): 4235-46.

Montini, E., D. Cesana, M. Schmidt, F. Sanvito, M. Ponzoni, C. Bartholomae, L. Sergi Sergi, F. Benedicenti, A. Ambrosi, C. Di Serio, C. Doglioni, C. von Kalle and L. Naldini (2006). "Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration." Nat Biotechnol **24**(6): 687-96.

Morgan, R. A., M. E. Dudley, J. R. Wunderlich, M. S. Hughes, J. C. Yang, R. M. Sherry, R. E. Royal, S. L. Topalian, U. S. Kammula, N. P. Restifo, Z. Zheng, A. Nahvi, C. R. de Vries, L. J. Rogers-Freezer, S. A. Mavroukakis and S. A. Rosenberg (2006). "Cancer regression in patients after transfer of genetically engineered lymphocytes." Science **314**(5796): 126-9.

Mortelmans, K. and E. Zeiger (2000). "The Ames Salmonella/microsome mutagenicity assay." Mutat Res **455**(1-2): 29-60.

Moses, A., J. Nelson and G. C. Bagby, Jr. (1998). "The influence of human immunodeficiency virus-1 on hematopoiesis." Blood **91**(5): 1479-95.

Mucenski, M. L., B. A. Taylor, J. N. Ihle, J. W. Hartley, H. C. Morse, 3rd, N. A. Jenkins and N. G. Copeland (1988). "Identification of a common ecotropic viral integration site, Evi-1, in the DNA of AKXD murine myeloid tumors." Mol Cell Biol **8**(1): 301-8.

Mui, A. L., H. Wakao, T. Kinoshita, T. Kitamura and A. Miyajima (1996). "Suppression of interleukin-3-induced gene expression by a C-terminal truncated Stat5: role of Stat5 in proliferation." Embo J **15**(10): 2425-33.

Mukhopadhyaya, R., J. Richardson, V. Nazarov, A. Corbin, R. Koller, M. Sitbon and L. Wolff (1994). "Different abilities of Friend murine leukemia virus (MuLV) and Moloney MuLV to induce promonocytic leukemia are due to determinants in both psi-gag-PR and env regions." J Virol **68**(8): 5100-7.

Nabel, G. and D. Baltimore (1987). "An inducible transcription factor activates expression of human immunodeficiency virus in T cells." Nature **326**(6114): 711-3.

Nakamura, T., R. Ouchida, T. Kodama, T. Kawashima, Y. Makino, N. Yoshikawa, S. Watanabe, C. Morimoto, T. Kitamura and H. Tanaka (2002). "Cytokine receptor common beta subunit-mediated STAT5 activation confers NF-kappa B activation in murine proB cell line Ba/F3 cells." J Biol Chem **277**(8): 6254-65.

Naldini, L., U. Blomer, P. Gallay, D. Ory, R. Mulligan, F. H. Gage, I. M. Verma and D. Trono (1996). "In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector." Science **272**(5259): 263-7.

Neil, S. J., T. Zang and P. D. Bieniasz (2008). "Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu." Nature **451**(7177): 425-30.

Nielsen, A. L., N. Pallisgaard, F. S. Pedersen and P. Jorgensen (1994). "Basic helix-loop-helix proteins in murine type C retrovirus transcriptional regulation." J Virol **68**(9): 5638-47.

Nosaka, T., T. Kawashima, K. Misawa, K. Ikuta, A. L. Mui and T. Kitamura (1999). "STAT5 as a molecular regulator of proliferation, differentiation and apoptosis in hematopoietic cells." Embo J **18**(17): 4754-65.

Onishi, M., A. L. Mui, Y. Morikawa, L. Cho, S. Kinoshita, G. P. Nolan, D. M. Gorman, A. Miyajima and T. Kitamura (1996). "Identification of an

oncogenic form of the thrombopoietin receptor MPL using retrovirus-mediated gene transfer." Blood **88**(4): 1399-406.

Onishi, M., T. Nosaka, K. Misawa, A. L. Mui, D. Gorman, M. McMahon, A. Miyajima and T. Kitamura (1998). "Identification and characterization of a constitutively active STAT5 mutant that promotes cell proliferation." Mol Cell Biol **18**(7): 3871-9.

Osborne, C. S., M. A. Vadas and P. N. Cockerill (1995). "Transcriptional regulation of mouse granulocyte-macrophage colony-stimulating factor/IL-3 locus." J Immunol **155**(1): 226-35.

Ott, M. G., M. Schmidt, K. Schwarzwaelder, S. Stein, U. Siler, U. Koehl, H. Glimm, K. Kuhlcke, A. Schilz, H. Kunkel, S. Naundorf, A. Brinkmann, A. Deichmann, M. Fischer, C. Ball, I. Pilz, C. Dunbar, Y. Du, N. A. Jenkins, N. G. Copeland, U. Luthi, M. Hassan, A. J. Thrasher, D. Hoelzer, C. von Kalle, R. Seger and M. Grez (2006). "Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1." Nat Med **12**(4): 401-9.

Ozaki, K., K. Kikly, D. Michalovich, P. R. Young and W. J. Leonard (2000). "Cloning of a type I cytokine receptor most related to the IL-2 receptor beta chain." Proc Natl Acad Sci U S A **97**(21): 11439-44.

Palacios, C., A. Gutierrez del Arroyo, A. Silva and M. K. Collins (2000). "The role of p53 in death of IL-3-dependent cells in response to cytotoxic drugs." Oncogene **19**(31): 3556-9.

Palacios, R., G. Henson, M. Steinmetz and J. P. McKearn (1984). "Interleukin-3 supports growth of mouse pre-B-cell clones in vitro." Nature **309**(5964): 126-31.

Palacios, R. and M. Steinmetz (1985). "Il-3-dependent mouse clones that express B-220 surface antigen, contain Ig genes in germ-line configuration, and generate B lymphocytes in vivo." Cell **41**(3): 727-34.

Palmer, E., C. Sporborg, A. Harrison, M. L. Martin and P. Feorino (1985). "Morphology and immunoelectron microscopy of AIDS virus." Arch Virol **85**(3-4): 189-96.

Panganiban, A. T. and H. M. Temin (1983). "The terminal nucleotides of retrovirus DNA are required for integration but not virus production." Nature **306**(5939): 155-60.

Patience, C., Y. Takeuchi, F. L. Cosset and R. A. Weiss (1998). "Packaging of endogenous retroviral sequences in retroviral vectors produced by murine and human packaging cells." J Virol **72**(4): 2671-6.

Paulus, W., I. Baur, F. M. Boyce, X. O. Breakefield and S. A. Reeves (1996). "Self-contained, tetracycline-regulated retroviral vector system for gene delivery to mammalian cells." J Virol **70**(1): 62-7.

Pelchen-Matthews, A., B. Kramer and M. Marsh (2003). "Infectious HIV-1 assembles in late endosomes in primary macrophages." J Cell Biol **162**(3): 443-55.

Perk, K. and J. E. Dahlberg (1974). "Murine intracisternal A type particles fail to separate from the membrane of the endoplasmic reticulum." J Virol **14**(5): 1304-6.

Perkins, N. D., A. B. Agranoff, C. S. Duckett and G. J. Nabel (1994). "Transcription factor AP-2 regulates human immunodeficiency virus type 1 gene expression." J Virol **68**(10): 6820-3.

Phillips, A. N. (1996). "Reduction of HIV concentration during acute infection: independence from a specific immune response." Science **271**(5248): 497-9.

Picker, L. J. (2006). "Immunopathogenesis of acute AIDS virus infection." Curr Opin Immunol **18**(4): 399-405.

Piguet, V. and Q. Sattentau (2004). "Dangerous liaisons at the virological synapse." J Clin Invest **114**(5): 605-10.

Piguet, V. and R. M. Steinman (2007). "The interaction of HIV with dendritic cells: outcomes and pathways." Trends Immunol **28**(11): 503-10.

Pless, M., K. Norga, M. Carroll, M. H. Heim, A. D. D'Andrea and B. Mathey-Prevot (1997). "Receptors that induce erythroid differentiation of Ba/F3 cells: structural requirements and effect on STAT5 binding." Blood **89**(9): 3175-85.

Pollok, K. E. and D. A. Williams (1999). "Facilitation of retrovirus-mediated gene transfer into hematopoietic stem and progenitor cells and peripheral blood T-lymphocytes utilizing recombinant fibronectin fragments." Curr Opin Mol Ther **1**(5): 595-604.

Prassolov, V., J. Meyer, G. Brandenburg, J. Hannemann, J. Bergemann, W. Ostertag and C. Stocking (2001). "Functional identification of secondary mutations inducing autonomous growth in synergy with a truncated interleukin-3 receptor: implications for multi-step oncogenesis." Exp Hematol **29**(6): 756-65.

Reddy, E. P., A. Korapati, P. Chaturvedi and S. Rane (2000). "IL-3 signaling and the role of Src kinases, JAKs and STATs: a covert liaison unveiled." Oncogene **19**(21): 2532-47.

Reisman, D. (1990). "Nuclear factor-1 (NF-1) binds to multiple sites within the transcriptional enhancer of Moloney murine leukemia virus." FEBS Lett **277**(1-2): 209-11.

Riviere, I., K. Brose and R. C. Mulligan (1995). "Effects of retroviral vector design on expression of human adenosine deaminase in murine bone marrow transplant recipients engrafted with genetically modified cells." Proc Natl Acad Sci U S A **92**(15): 6733-7.

Robb, L. (2007). "Cytokine receptors and hematopoietic differentiation." Oncogene **26**(47): 6715-23.

Rodriguez-Tarduchy, G., M. K. Collins, I. Garcia and A. Lopez-Rivas (1992). "Insulin-like growth factor-I inhibits apoptosis in IL-3-dependent hemopoietic cells." J Immunol **149**(2): 535-40.

Rosas, M., S. Gordon and P. R. Taylor (2007). "Characterisation of the expression and function of the GM-CSF receptor alpha-chain in mice." Eur J Immunol **37**(9): 2518-28.

Rosenberg, S. A., P. Aebersold, K. Cornetta, A. Kasid, R. A. Morgan, R. Moen, E. M. Karson, M. T. Lotze, J. C. Yang, S. L. Topalian and et al. (1990). "Gene transfer into humans--immunotherapy of patients with advanced melanoma, using tumor-infiltrating lymphocytes modified by retroviral gene transduction." N Engl J Med **323**(9): 570-8.

Rossi, J. J., C. H. June and D. B. Kohn (2007). "Genetic therapies against HIV." Nat Biotechnol **25**(12): 1444-54.

Rous, P. (1979). "A transmissible avian neoplasm. (Sarcoma of the common fowl) by Peyton Rous, M.D., Experimental Medicine for Sept. 1, 1910, vol. 12, pp.696-705." J Exp Med **150**(4): 738-53.

Sanders, D. A. (2002). "No false start for novel pseudotyped vectors." Curr Opin Biotechnol **13**(5): 437-42.

Sasaki, M., T. Sasaki and J. Enami (1996). "Prolactin-dependent growth and gamma-casein gene expression in Ba/F3 cells transfected with a long form of mouse mammary prolactin receptor." Endocr J **43**(1): 45-52.

Schambach, A., M. Galla, T. Maetzig, R. Loew and C. Baum (2007). "Improving transcriptional termination of self-inactivating gamma-retroviral and lentiviral vectors." Mol Ther **15**(6): 1167-73.

Scheid, M. P., K. M. Schubert and V. Duronio (1999). "Regulation of bad phosphorylation and association with Bcl-x(L) by the MAPK/Erk kinase." J Biol Chem **274**(43): 31108-13.

Schiff, R. D. and A. Oliff (1986). "The pathophysiology of murine retrovirus-induced leukemias." Crit Rev Oncol Hematol **5**(3): 257-323.

Schroder, A. R., P. Shinn, H. Chen, C. Berry, J. R. Ecker and F. Bushman (2002). "HIV-1 integration in the human genome favors active genes and local hotspots." Cell **110**(4): 521-9.

Schwartzbauer, G., J. H. Yu, H. Cheng and R. K. Menon (1998). "Transcription factor MSY-1 regulates expression of the murine growth hormone receptor gene." J Biol Chem **273**(38): 24760-9.

Schwarzwaelder, K., S. J. Howe, M. Schmidt, M. H. Brugman, A. Deichmann, H. Glimm, S. Schmidt, C. Prinz, M. Wissler, D. J. King, F. Zhang, K. L. Parsley, K. C. Gilmour, J. Sinclair, J. Bayford, R. Peraj, K. Pike-Overzet, F. J. Staal, D. de Ridder, C. Kinnon, U. Abel, G. Wagemaker, H. B. Gaspar, A. J. Thrasher and C. von Kalle (2007). "Gammaretrovirus-mediated correction of SCID-X1 is associated with skewed vector integration site distribution in vivo." J Clin Invest **117**(8): 2241-9.

Seggewiss, R., S. Pittaluga, R. L. Adler, F. J. Guenaga, C. Ferguson, I. H. Pilz, B. Ryu, B. P. Sorrentino, W. S. Young, 3rd, R. E. Donahue, C. von Kalle, A. W. Nienhuis and C. E. Dunbar (2006). "Acute myeloid leukemia is associated with retroviral gene transfer to hematopoietic progenitor cells in a rhesus macaque." Blood **107**(10): 3865-7.

Selten, G., H. T. Cuypers and A. Berns (1985). "Proviral activation of the putative oncogene Pim-1 in MuLV induced T-cell lymphomas." Embo J **4**(7): 1793-8.

Serwadda, D., R. D. Mugerwa, N. K. Sewankambo, A. Lwegaba, J. W. Carswell, G. B. Kirya, A. C. Bayley, R. G. Downing, R. S. Tedder, S. A. Clayden and et al. (1985). "Slim disease: a new disease in Uganda and its association with HTLV-III infection." Lancet **2**(8460): 849-52.

Sheehy, A. M., N. C. Gaddis and M. H. Malim (2003). "The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif." Nat Med **9**(11): 1404-7.

Shehu-Xhilaga, M., S. M. Crowe and J. Mak (2001). "Maintenance of the Gag/Gag-Pol ratio is important for human immunodeficiency virus type 1 RNA dimerization and viral infectivity." J Virol **75**(4): 1834-41.

Shen-Ong, G. L., H. C. Morse, 3rd, M. Potter and J. F. Mushinski (1986). "Two modes of c-myb activation in virus-induced mouse myeloid tumors." Mol Cell Biol **6**(2): 380-92.

Shen-Ong, G. L. and L. Wolff (1987). "Moloney murine leukemia virus-induced myeloid tumors in adult BALB/c mice: requirement of c-myb activation but lack of v-abl involvement." J Virol **61**(12): 3721-5.

Shiramizu, B., B. G. Herndier and M. S. McGrath (1994). "Identification of a common clonal human immunodeficiency virus integration site in human immunodeficiency virus-associated lymphomas." Cancer Res **54**(8): 2069-72.

Shou, Y., Z. Ma, T. Lu and B. P. Sorrentino (2006). "Unique risk factors for insertional mutagenesis in a mouse model of XSCID gene therapy." Proc Natl Acad Sci U S A **103**(31): 11730-5.

Silver, J. and V. Keerikatte (1989). "Novel use of polymerase chain reaction to amplify cellular DNA adjacent to an integrated provirus." J Virol **63**(5): 1924-8.

Simon, V. and D. D. Ho (2003). "HIV-1 dynamics in vivo: implications for therapy." Nat Rev Microbiol **1**(3): 181-90.

Smagulova, F., S. Maurel, Z. Morichaud, C. Devaux, M. Mougel and L. Houzet (2005). "The highly structured encapsidation signal of MuLV RNA is involved in the nuclear export of its unspliced RNA." J Mol Biol **354**(5): 1118-28.

Soneoka, Y., P. M. Cannon, E. E. Ramsdale, J. C. Griffiths, G. Romano, S. M. Kingsman and A. J. Kingsman (1995). "A transient three-plasmid expression system for the production of high titer retroviral vectors." Nucleic Acids Res **23**(4): 628-33.

Speck, N. A., B. Renjifo, E. Golemis, T. N. Fredrickson, J. W. Hartley and N. Hopkins (1990). "Mutation of the core or adjacent LVb elements of the Moloney murine leukemia virus enhancer alters disease specificity." Genes Dev **4**(2): 233-42.

Speck, N. A., B. Renjifo and N. Hopkins (1990). "Point mutations in the Moloney murine leukemia virus enhancer identify a lymphoid-specific viral core motif and 1,3-phorbol myristate acetate-inducible element." J Virol **64**(2): 543-50.

Steffens, C. M. and T. J. Hope (2001). "Recent advances in the understanding of HIV accessory protein function." Aids **15 Suppl 5**: S21-6.

Stehelin, D., H. E. Varmus, J. M. Bishop and P. K. Vogt (1976). "DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA." Nature **260**(5547): 170-3.

Stevens, M., E. De Clercq and J. Balzarini (2006). "The regulation of HIV-1 transcription: molecular targets for chemotherapeutic intervention." Med Res Rev **26**(5): 595-625.

Stewart, M., N. Mackay, L. Hanlon, K. Blyth, L. Scobie, E. Cameron and J. C. Neil (2007). "Insertional mutagenesis reveals progression genes and checkpoints in MYC/Runx2 lymphomas." Cancer Res **67**(11): 5126-33.

Stocking, C., U. Bergholz, J. Friel, K. Klingler, T. Wagener, C. Starke, T. Kitamura, A. Miyajima and W. Ostertag (1993). "Distinct classes of factor-independent mutants can be isolated after retroviral mutagenesis of a human myeloid stem cell line." Growth Factors **8**(3): 197-209.

Stocking, C., C. Loliger, M. Kawai, S. Suciu, N. Gough and W. Ostertag (1988). "Identification of genes involved in growth autonomy of hematopoietic cells by analysis of factor-independent mutants." Cell **53**(6): 869-79.

Suzuki, T., K. Minehata, K. Akagi, N. A. Jenkins and N. G. Copeland (2006). "Tumor suppressor gene identification using retroviral insertional mutagenesis in Blm-deficient mice." Embo J **25**(14): 3422-31.

Suzuki, T., H. Shen, K. Akagi, H. C. Morse, J. D. Malley, D. Q. Naiman, N. A. Jenkins and N. G. Copeland (2002). "New genes involved in cancer identified by retroviral tagging." Nat Genet **32**(1): 166-74.

Suzuki, Y. and R. Craigie (2007). "The road to chromatin - nuclear entry of retroviruses." Nat Rev Microbiol **5**(3): 187-96.

Suzuki, Y., H. Yang and R. Craigie (2004). "LAP2alpha and BAF collaborate to organize the Moloney murine leukemia virus preintegration complex." Embo J **23**(23): 4670-8.

Tailor, C. S., A. Nouri, C. G. Lee, C. Kozak and D. Kabat (1999). "Cloning and characterization of a cell surface receptor for xenotropic and polytropic murine leukemia viruses." Proc Natl Acad Sci U S A **96**(3): 927-32.

Takeuchi, Y. and M. Pizzato (2000). "Retrovirus vectors." Adv Exp Med Biol **465**: 23-35.

Takeuchi, Y., C. D. Porter, K. M. Strahan, A. F. Preece, K. Gustafsson, F. L. Cosset, R. A. Weiss and M. K. Collins (1996). "Sensitization of cells and retroviruses to human serum by (alpha 1-3) galactosyltransferase." Nature **379**(6560): 85-8.

Temin, H. M. and S. Mizutani (1970). "RNA-dependent DNA polymerase in virions of Rous sarcoma virus." Nature **226**(5252): 1211-3.

Themis, M., D. May, C. Coutelle and R. F. Newbold (2003). "Mutational effects of retrovirus insertion on the genome of V79 cells by an attenuated retrovirus vector: implications for gene therapy." Gene Ther **10**(19): 1703-11.

Themis, M., S. N. Waddington, M. Schmidt, C. von Kalle, Y. Wang, F. Al-Allaf, L. G. Gregory, M. Nivsarkar, M. Themis, M. V. Holder, S. M. Buckley, N. Dighe, A. T. Ruthe, A. Mistry, B. Bigger, A. Rahim, T. H. Nguyen, D. Trono, A. J. Thrasher and C. Coutelle (2005). "Oncogenesis following delivery of a nonprimate lentiviral gene therapy vector to fetal and neonatal mice." Mol Ther **12**(4): 763-71.

Thomas, C. E., A. Ehrhardt and M. A. Kay (2003). "Progress and problems with the use of viral vectors for gene therapy." Nat Rev Genet **4**(5): 346-58.

Thomas, J., Y. Leverrier and J. Marvel (1998). "Bcl-X is the major pleiotropic anti-apoptotic gene activated by retroviral insertion mutagenesis in an IL-3 dependent bone marrow derived cell line." Oncogene **16**(11): 1399-408.

Thomas, J., Y. Leverrier, A. L. Mathieu and J. Marvel (2004). "Identification of apoptosis regulatory genes using insertional mutagenesis." Methods Mol Biol **282**: 275-89.

Thrasher, A. J., H. B. Gaspar, C. Baum, U. Modlich, A. Schambach, F. Candotti, M. Otsu, B. Sorrentino, L. Scobie, E. Cameron, K. Blyth, J. Neil, S. H. Abina, M. Cavazzana-Calvo and A. Fischer (2006). "Gene therapy: X-SCID transgene leukaemogenicity." Nature **443**(7109): E5-6; discussion E-7.

Towatari, M., Y. Kanei, H. Saito and M. Hamaguchi (1998). "Hematopoietic transcription factor GATA-2 activates transcription from HIV-1 long terminal repeat." Aids **12**(3): 253-9.

Towatari, M., G. E. May, R. Marais, G. R. Perkins, C. J. Marshall, S. Cowley and T. Enver (1995). "Regulation of GATA-2 phosphorylation by mitogen-activated protein kinase and interleukin-3." J Biol Chem **270**(8): 4101-7.

Uren, A. G., J. Kool, A. Berns and M. van Lohuizen (2005). "Retroviral insertional mutagenesis: past, present and future." Oncogene **24**(52): 7656-72.

Valsamakis, A., S. Zeichner, S. Carswell and J. C. Alwine (1991). "The human immunodeficiency virus type 1 polyadenylylation signal: a 3' long terminal repeat element upstream of the AAUAAA necessary for efficient polyadenylylation." Proc Natl Acad Sci U S A **88**(6): 2108-12.

Van Lint, C., S. Emiliani, M. Ott and E. Verdin (1996). "Transcriptional activation and chromatin remodeling of the HIV-1 promoter in response to histone acetylation." Embo J **15**(5): 1112-20.

van Lohuizen, M., M. Breuer and A. Berns (1989). "N-myc is frequently activated by proviral insertion in MuLV-induced T cell lymphomas." Embo J **8**(1): 133-6.

van Zeijl, M., S. V. Johann, E. Closs, J. Cunningham, R. Eddy, T. B. Shows and B. O'Hara (1994). "A human amphotropic retrovirus receptor is a second member of the gibbon ape leukemia virus receptor family." Proc Natl Acad Sci U S A **91**(3): 1168-72.

Varmus, H. E., N. Quintrell and S. Ortiz (1981). "Retroviruses as mutagens: insertion and excision of a nontransforming provirus alter expression of a resident transforming provirus." Cell **25**(1): 23-36.

Verdin, E., P. Paras, Jr. and C. Van Lint (1993). "Chromatin disruption in the promoter of human immunodeficiency virus type 1 during transcriptional activation." Embo J **12**(8): 3249-59.

Verhoeyen, E. and F. L. Cosset (2004). "Surface-engineering of lentiviral vectors." J Gene Med **6 Suppl 1**: S83-94.

Wang, C. L., B. B. Wang, G. Bartha, L. Li, N. Channa, M. Klinger, N. Killeen and M. Wabl (2006). "Activation of an oncogenic microRNA cistron by provirus integration." Proc Natl Acad Sci U S A **103**(49): 18680-4.

Wang, J. K., G. Gao and M. Goldfarb (1994). "Fibroblast growth factor receptors have different signaling and mitogenic potentials." Mol Cell Biol **14**(1): 181-8.

Wang, S., Q. Wang, B. E. Crute, I. N. Melnikova, S. R. Keller and N. A. Speck (1993). "Cloning and characterization of subunits of the T-cell receptor and murine leukemia virus enhancer core-binding factor." Mol Cell Biol **13**(6): 3324-39.

Wang, X. Y., L. S. Steelman and J. A. McCubrey (1997). "Abnormal activation of cytokine gene expression by intracisternal type A particle transposition: effects of mutations that result in autocrine growth stimulation and malignant transformation." Cytokines Cell Mol Ther **3**(1): 3-19.

Watanabe, S., T. Murakami, T. Nakamura, C. Morimoto and K. Arai (2002). "Human GM-CSF induces HIV-1 LTR by multiple signalling pathways." Biochimie **84**(7): 633-42.

Wei, P., M. E. Garber, S. M. Fang, W. H. Fischer and K. A. Jones (1998). "A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA." Cell **92**(4): 451-62.

West, M. J., A. D. Lowe and J. Karn (2001). "Activation of human immunodeficiency virus transcription in T cells revisited: NF-kappaB p65 stimulates transcriptional elongation." J Virol **75**(18): 8524-37.

Williams, K. J. and L. A. Loeb (1992). "Retroviral reverse transcriptases: error frequencies and mutagenesis." Curr Top Microbiol Immunol **176**: 165-80.

Williams, S. A., L. F. Chen, H. Kwon, C. M. Ruiz-Jarabo, E. Verdin and W. C. Greene (2006). "NF-kappaB p50 promotes HIV latency through HDAC recruitment and repression of transcriptional initiation." Embo J **25**(1): 139-49.

Williams, S. A. and W. C. Greene (2007). "Regulation of HIV-1 latency by T-cell activation." Cytokine **39**(1): 63-74.

Williams, S. A., H. Kwon, L. F. Chen and W. C. Greene (2007). "Sustained induction of NF-kappa B is required for efficient expression of latent human immunodeficiency virus type 1." J Virol **81**(11): 6043-56.

Wolf, D. and S. P. Goff (2007). "TRIM28 Mediates Primer Binding Site-Targeted Silencing of Murine Leukemia Virus in Embryonic Cells." Cell **131**(1): 46-57.

Wolff, L. and R. Koller (1990). "Regions of the Moloney murine leukemia virus genome specifically related to induction of promonocytic tumors." J Virol **64**(1): 155-60.

Woods, N. B., V. Bottero, M. Schmidt, C. von Kalle and I. M. Verma (2006). "Gene therapy: therapeutic gene causing lymphoma." Nature **440**(7088): 1123.

Woods, N. B., A. Ooka and S. Karlsson (2002). "Development of gene therapy for hematopoietic stem cells using lentiviral vectors." Leukemia **16**(4): 563-9.

Wool-Lewis, R. J. and P. Bates (1998). "Characterization of Ebola virus entry by using pseudotyped viruses: identification of receptor-deficient cell lines." J Virol **72**(4): 3155-60.

Wu, X., Y. Li, B. Crise and S. M. Burgess (2003). "Transcription start regions in the human genome are favored targets for MLV integration." Science **300**(5626): 1749-51.

Yamada, G., Y. Kitamura, H. Sonoda, H. Harada, S. Taki, R. C. Mulligan, H. Osawa, T. Diamantstein, S. Yokoyama and T. Taniguchi (1987). "Retroviral expression of the human IL-2 gene in a murine T cell line results in cell growth autonomy and tumorigenicity." Embo J **6**(9): 2705-9.

Yamashita, M. and M. Emerman (2004). "Capsid is a dominant determinant of retrovirus infectivity in nondividing cells." J Virol **78**(11): 5670-8.

Yamashita, M. and M. Emerman (2005). "The cell cycle independence of HIV infections is not determined by known karyophilic viral elements." PLoS Pathog **1**(3): e18.

Yamashita, M., O. Perez, T. J. Hope and M. Emerman (2007). "Evidence for direct involvement of the capsid protein in HIV infection of nondividing cells." PLoS Pathog **3**(10): 1502-10.

Yang, L., L. Bailey, D. Baltimore and P. Wang (2006). "Targeting lentiviral vectors to specific cell types in vivo." Proc Natl Acad Sci U S A **103**(31): 11479-84.

Yang, L., H. Yang, K. Rideout, T. Cho, K. I. Joo, L. Ziegler, A. Elliot, A. Walls, D. Yu, D. Baltimore and P. Wang (2008). "Engineered lentivector targeting of dendritic cells for in vivo immunization." Nat Biotechnol **26**(3): 326-34.

Yang, X., Y. Chen and D. Gabuzda (1999). "ERK MAP kinase links cytokine signals to activation of latent HIV-1 infection by stimulating a cooperative interaction of AP-1 and NF-kappaB." J Biol Chem **274**(39): 27981-8.

Yang, Y. L., L. Guo, S. Xu, C. A. Holland, T. Kitamura, K. Hunter and J. M. Cunningham (1999). "Receptors for polytropic and xenotropic mouse leukaemia viruses encoded by a single gene at Rmc1." Nat Genet **21**(2): 216-9.

Yee, J. K., J. C. Moores, D. J. Jolly, J. A. Wolff, J. G. Respess and T. Friedmann (1987). "Gene expression from transcriptionally disabled retroviral vectors." Proc Natl Acad Sci U S A **84**(15): 5197-201.

Ymer, S., W. Q. Tucker, H. D. Campbell and I. G. Young (1986). "Nucleotide sequence of the intracisternal A-particle genome inserted 5' to the interleukin-3 gene of the leukemia cell line WEHI-3B." Nucleic Acids Res **14**(14): 5901-18.

Ymer, S., W. Q. Tucker, C. J. Sanderson, A. J. Hapel, H. D. Campbell and I. G. Young (1985). "Constitutive synthesis of interleukin-3 by leukaemia cell line WEHI-3B is due to retroviral insertion near the gene." Nature **317**(6034): 255-8.

Yoshizato, H., M. Tanaka, T. Fujikawa, Y. Higashimoto, A. Shimizu and K. Nakashima (2000). "20 kDa human growth hormone (20K hGH) stimulates insulin-like growth factor-I (IGF-I) gene expression at lower concentrations than 22K hGH in hGH receptor-expressing Ba/F3 cells." Endocr J **47 Suppl**: S37-40.

Yu, J. H., G. Schwartzbauer, A. Kazlman and R. K. Menon (1999). "Role of the Sp family of transcription factors in the ontogeny of growth hormone receptor gene expression." J Biol Chem **274**(48): 34327-36.

Yu, S. F., T. von Ruden, P. W. Kantoff, C. Garber, M. Seiberg, U. Ruther, W. F. Anderson, E. F. Wagner and E. Gilboa (1986). "Self-inactivating retroviral vectors designed for transfer of whole genes into mammalian cells." Proc Natl Acad Sci U S A **83**(10): 3194-8.

Yu, Y. L., Y. J. Chiang, Y. C. Chen, M. Papetti, C. G. Juo, A. I. Skoultchi and J. J. Yen (2005). "MAPK-mediated phosphorylation of GATA-1 promotes Bcl-XL expression and cell survival." J Biol Chem **280**(33): 29533-42.

Yu, Y. L., Y. J. Chiang and J. J. Yen (2002). "GATA factors are essential for transcription of the survival gene E4bp4 and the viability response of interleukin-3 in Ba/F3 hematopoietic cells." J Biol Chem **277**(30): 27144-53.

Zaitseva, L., R. Myers and A. Fassati (2006). "tRNAs promote nuclear import of HIV-1 intracellular reverse transcription complexes." PLoS Biol **4**(10): e332.

Zennou, V., C. Petit, D. Guetard, U. Nerhbass, L. Montagnier and P. Charneau (2000). "HIV-1 genome nuclear import is mediated by a central DNA flap." Cell **101**(2): 173-85.

Zhong, H., M. J. May, E. Jimi and S. Ghosh (2002). "The phosphorylation status of nuclear NF-kappa B determines its association with CBP/p300 or HDAC-1." Mol Cell **9**(3): 625-36.

Zhou, M., M. A. Halanski, M. F. Radonovich, F. Kashanchi, J. Peng, D. H. Price and J. N. Brady (2000). "Tat modifies the activity of CDK9 to phosphorylate serine 5 of the RNA polymerase II carboxyl-terminal domain during human immunodeficiency virus type 1 transcription." Mol Cell Biol **20**(14): 5077-86.

Zufferey, R., J. E. Donello, D. Trono and T. J. Hope (1999). "Woodchuck hepatitis virus posttranscriptional regulatory element enhances expression of transgenes delivered by retroviral vectors." J Virol **73**(4): 2886-92.

Zufferey, R., T. Dull, R. J. Mandel, A. Bukovsky, D. Quiroz, L. Naldini and D. Trono (1998). "Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery." J Virol **72**(12): 9873-80.

Zufferey, R., D. Nagy, R. J. Mandel, L. Naldini and D. Trono (1997). "Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo." Nat Biotechnol **15**(9): 871-5.

Zychlinski, D., A. Schambach, U. Modlich, T. Maetzig, J. Meyer, E. Grassman, A. Mishra and C. Baum (2008). "Physiological promoters reduce the genotoxic risk of integrating gene vectors." Mol Ther **16**(4): 718-25.